V.A. AMBARTSUMIAN

THEORETICAL ASTROPHYSICS

V.A. AMBARTSUMIAN

THEORETICAL ASTROPHYSICS

Contents

Part I. The theory of the radiative equilibrium of stellar photospheres and the continuous spectrum of stars

Part VIII. The scattering of light in planetary atmospheres

Part IX. Interstellar matter

FROM THE PREFACE TO THE RUSSIAN EDITION

THEORETICAL astrophysics is the science which studies and explains, using the laws of physics, the physical processes occurring in the heavenly bodies. In doing so, wide use is made of mathematical methods, but these play only a subsidiary part.

Although theoretical astrophysics is a young and very rapidly developing science, its results are already of great importance in all branches of astronomy and in many branches of physics. The subject has been developed particularly in the USSR.

The first Russian textbook on theoretical astrophysics appeared in 1939, and was written by the editor of the present work. During the past twelve years, this young branch of astronomy has made so much progress and has undergone such profound changes that it was necessary to abandon the idea of merely revising the old book, and to compile an entirely new one.

In 1939 the main content of theoretical astrophysics was formed by problems of the radiative transfer of energy and of the excitation of atoms in stellar atmospheres and in the nebulae. Today, although the treatment of the radiation field and radiative equilibrium has been considerably developed and refined, the processes occuring in the macroscopic electromagnetic fields which exist in the stars and in the Sun are becoming more and more important in theoretical astrophysics. It is easy to see why these problems began to be developed much later than those of the theory of radiative equilibrium. By observing the radiation of the stars and nebulae, direct information is obtained about the radiation field in the form of values for the intensity of radiation emerging from these bodies. These results are both the basis and the test of the theory of radiative equilibrium. The study of the macroscopic electromagnetic fields is more involved, since we must draw our conclusions from the effects of the fields on the motions of charged particles. The observation of these motions is very difficult, and is possible only by the use of new techniques. The importance of these fields is fairly well understood only in the case of the Sun.

Soviet scientists who work in the field of theoretical astrophysics are guided by the method of dialectical materialism, and always link their work to reality, using mathematical and physical methods as an important and powerful means of studying the heavenly bodies.

In the field of theoretical astrophysics, Soviet scientists now occupy the leading place in the world. The authors of this book have made by their investigations considerable contributions to the branches of theoretical astrophysics which they discuss. The treatment of many topics in the present work is therefore completely original, and sometimes differs sharply from anything that has been published in scientific literature. This fact should be kept in mind in studying the book.

Parts I, II, and III were written by É. R. MUSTEL', Parts IV, V, and VI by V. V. SOBOLEV, Part VII by A. B. SEVERNYĬ, and Parts VIII and IX by V. A. AMBARTSUMYAN.

The authors express their gratitude to S. B. PIKEL'NER, who wrote Chapters 18 and 22 in Part III.

The course is designed for university students, graduates, and scientific workers. In a new branch of science, errors are inevitable, but it is hoped that this book contains as few as possible.

We shall be very grateful to any reader who points out errors or submits critical remarks.

1952 V. A. AMBARTSUMYAN

PREFACE TO THE ENGLISH EDITION

FOUR years have passed since the Russian edition of this book was published. This is no little time in the rapidly developing subject of astrophysics. When the Pergamon Press undertook to publish an English edition, they asked me to examine the MS of the translation and to make any necessary changes and additions. This has been done by my fellow authors and myself. We have confined ourselves, however, to the most essential additions and corrections, and have not rewritten any part of the book.

The pioneers of theoretical astrophysics — JEANS, EDDINGTON, MILNE — were Englishmen, and wrote in English. The researches of the present generation of British scientists in this field are also very well known. We, the authors, therefore, while happy to see our book appear in English, have also a feeling of responsibility.

We are grateful to the Pergamon Press for their initiative and perseverance, and to the translator for what is, so far as we can judge, an exemplary piece of work.

May 1956

V. A. AMBARTSUMYAN
Editor of the Russian edition

NOTE

THE system of transliteration used is that recommended by the Royal Society. All names of Russian astronomers have been rendered in accordance with this system, but the Index of Names contains many of the alternative forms that have been used at various times.

References to Russian journals have been retained, but those to Russian textbooks not at present available in translation have been omitted.

My thanks are due to numerous colleagues for their valuable advice, and to Miss HEATHER CARTER for her labours at the typewriter.

J. B. S.

PART I.
THE THEORY OF THE RADIATIVE EQUILIBRIUM OF STELLAR PHOTOSPHERES AND THE CONTINUOUS SPECTRUM OF STARS

Chapter 1. Introductory remarks

IN Part I of this book we shall consider questions connected with the physical structure of stellar photospheres. Observations of the nearest star to us, the Sun, have enabled us to establish that its atmosphere may be divided into a series of layers differing from one another in their physical characteristics. The lowest and densest part of the solar atmosphere is called the **photosphere.** The photospheric layers emit practically all the radiant energy which the Sun sends out into space. The *continuous spectrum* of the Sun is basically the spectrum of the photospheric radiation.

The absorption lines in the solar spectrum (apart from the centres of strong absorption lines) are formed in approximately the same layers of the photosphere as the continuous spectrum. The division of the lower layers of the Sun's atmosphere into photosphere and reversing layer, which is frequently made in the literature, is of an arbitrary character.

Above the photosphere of the Sun lie the **chromosphere** (thickness about 15,000 km) and the **solar corona** (whose extent, including coronal rays, amounts to several solar radii). The chromosphere and corona play a negligible part in the formation of the continuous spectrum of the Sun in the visible region, and are important only in the far ultra-violet region of the spectrum (beyond the limit of the Lyman series) and in the region of very long waves (including the ultra-short radio-wave range).

As yet we know very little of the existence of a chromosphere or corona in stars which are essentially different from the Sun as regards spectral type. The existence of photospheric layers, however, is a necessary property of every star, for it is these layers which determine its luminosity. The interrelation between the photospheric layers and the layers in which the absorption lines are produced is a more complex question. A number of physical considerations show that, for the majority of stars of non-solar type, these layers are more or less identical also.

Thus the general theoretical considerations which we shall discuss in Part I of this book are applicable not only to the Sun, but also to a very large group of stars. However, in selecting objects for study we shall make a series of restrictions. Firstly, we shall study only stars whose luminosity does not vary with time, or varies only slowly. Secondly, we shall limit our investigations to stars where the thickness of the photospheric layers is very small in comparison with the radius. Finally, we exclude from consideration stars characterised by some peculiarity of spectrum (Wolf-Rayet stars, P Cygni type stars, etc.).

Let us consider these limitations in more detail. In stars of constant luminosity, the state of the photosphere may be regarded as stable and invariable in time. Of course, this statement is valid only *on the average*. In reality, as observation shows, the different parts of the surface of such a "constant" star as the Sun, for instance, undergo considerable changes at times (sunspots, faculae, etc.) Yet we may suppose that the state of the photospheric layers of the Sun as a whole is invariant. The same is true also of other stars of constant luminosity.

We turn now to our second restriction. The study of the Sun has shown that the thickness of its photosphere is very small and amounts to about 100 to 300 km. In comparison with the Sun's radius (696,300 km) this is, of course, an extremely small quantity. Hence, in our problems, the solar photosphere may be considered as composed of plane-parallel layers. Theoretical calculations show that this holds also for the majority of other non-variable stars.

Finally, Wolf-Rayet stars, P Cygni type stars, etc., form a very special group of stars; we shall consider them later.

It is to be noticed that the limitations which we have introduced exclude from consideration only a negligible proportion of the stars. Hence our theoretical results will be valid for a very large group of objects.

The main problems which the theory of stellar photospheres proposes to solve are essentially as follows:

(1) To find the law of variation of temperature, pressure, density, and other physical characteristics with depth in the stellar (or solar) atmosphere.

(2) To explain the properties of the continuous spectra of stars and of the Sun.

(3) To examine the law of variation of brightness over the discs of the Sun and of the stars.

The first of these problems is the most important; its solution determines that of the other two.

Chapter 2. Basic concepts of the theory of radiation. The equations of transfer

1. **The part played by radiation in stellar photospheres.** The physical state of any element of matter inside a stellar photosphere is determined by the interaction of this element with the surrounding medium. The chief question in which we are interested is by what means the transfer of energy takes place in the photospheric layers. In this connection we shall speak only of thermal energy. *In the term "thermal energy" we include the internal kinetic energy of the thermal motion of particles, the excitation energy of atoms and the ionisation energy* (see below).

The remaining forms of internal energy are unimportant in the photospheres of ordinary stars. For instance, nuclear processes (with release of energy) do not play a noticeable part in stellar photospheres. The following processes must be considered as being capable of bringing about the *exchange* of thermal energy in a stellar photosphere.

(1) The transfer of thermal energy by *conduction.*

(2) The transfer of thermal energy by *convection*, i. e. the transfer of heat by the actual movement of masses of gas.

(3) The transfer of thermal energy by *radiation.* Here it is understood that the radiation itself is of purely thermal character, i. e. is determined only by the temperature of the gas (which, of course, changes from point to point).

Numerous investigations have shown that the first process is negligible in stellar photospheres. These investigations have also established that, in the overwhelming majority of cases, at least for the stars to which we restrict ourselves, the exchange of thermal energy takes place mainly by thermal radiation; we shall see this from later considerations. We therefore now turn our main attention to questions relating to *thermal radiation*, and first recall some basic ideas of the theory of thermal radiation.

2. **The intensity and flux of radiation.** We consider a cavity penetrated in all directions by radiation. In this cavity we select an arbitrarily oriented small area $d\sigma$ and erect on it at the point P the normal n (Fig. 1). Next, at an angle θ to the normal we draw a line L, which we take as the axis of an elementary cone of solid angle $d\omega$. If we draw through every point of the boundary of the area $d\sigma$ a line parallel to the nearest generator of the cone

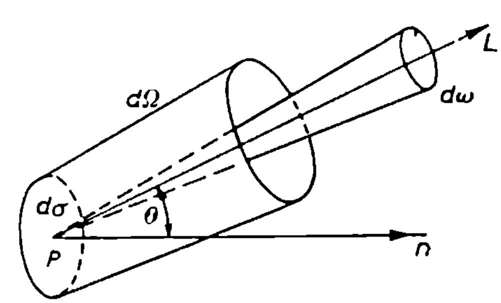

FIG. 1

1*

$d\omega$, then we shall have a truncated semi-infinite cone $d\Omega$, similar to the cone $d\omega$. Its cross-sectional area, perpendicular to L, at the point P will be $d\sigma \cos\theta$.

Let dE_ν be the total quantity of energy, passing in time dt through the area $d\sigma$ inside† the cone $d\Omega$, which belongs to the frequency interval between ν and $\nu + d\nu$. Then the specific **intensity** of **radiation**, or simply the **intensity**, I_ν, is defined as the following limit:

$$I_\nu = \lim_{d\sigma,\, dt,\, d\omega,\, d\nu \to 0} \frac{dE_\nu}{d\sigma \cos\theta \, dt \, d\omega \, d\nu} . \tag{2.1}$$

This limit is in general a function of the co-ordinates (i. e. of the position in space) of the point P, of the direction L, of the time t, and of the frequency ν.

The appearance of $\cos\theta$ in (2.1) is explained by the fact that we are considering a beam going not in the direction of the normal n but in the direction L; the quantity of energy which travels inside the solid angle $d\Omega$ is determined not by the area $d\sigma$ itself, but by its projection on a plane perpendicular to the direction L.

From the definition (2.1), we can calculate the quantity of energy dE_ν if we know I_ν:

$$dE_\nu = I_\nu \, d\sigma \cos\theta \, dt \, d\omega \, d\nu . \tag{2.2}$$

For the quantities dE and I, which are integrals over the whole spectrum, we have, instead of (2.2),

$$dE = I \, d\sigma \cos\theta \, dt \, d\omega . \tag{2.3}$$

The quantity of radiant energy dE_ν in (2.2) relates to the solid angle $d\omega$. The total quantity of energy of radiation dE^*_ν passing through the area $d\sigma$ in *all* directions will be:

$$dE^*_\nu = d\sigma \, dt \, d\nu \int_{4\pi} I_\nu \cos\theta \, d\omega . \tag{2.4}$$

Since for $\theta > \frac{1}{2}\pi$ the factor $\cos\theta$ is negative, dE^*_ν is in fact the excess of energy passing outwards through the area $d\sigma$ in time dt and frequency interval $d\nu$, over the energy passing inwards through that area. The magnitude of this excess for unit time, referred to *unit* area

† That is, along those directions which after intersecting the area $d\sigma$ remain entirely within $d\Omega$. We obtain the aggregate of such rays if we construct at every point of the area $d\sigma$ a cone containing all directions parallel to those contained in $d\omega$. It is clear from this that, for small $d\sigma$ and $d\omega$, the energy passing through $d\sigma$ inside $d\Omega$ will be proportional to $d\sigma \, d\omega$.

and *unit* frequency interval, is called the **flux of radiation** πH_ν (the factor π is introduced for convenience in later calculations). The flux πH_ν is obtained on dividing dE_ν^* by $d\sigma\, dt\, d\nu$, and thus

$$\pi H_\nu = \int_{4\pi} I_\nu \cos\theta\, d\omega\,. \tag{2.5}$$

The total (integrated) flux πH is clearly

$$\pi H = \pi \int_0^\infty H_\nu\, d\nu = \int_0^\infty \int_{4\pi} I_\nu \cos\theta\, d\omega\, d\nu\,. \tag{2.6}$$

It has the dimensions of energy per square centimetre per second. The flux quantities πH_ν and πH are, in general, functions of the space co-ordinates and orientation of the area.

We now introduce the expression for the solid angle $d\omega$. To do this, it is most convenient to use a system of spherical co-ordinates with its origin at the given point in the stellar atmosphere and its polar axis along the radius of the star. Since, by definition, the solid angle is numerically equal to the area intercepted on a sphere of unit radius by the corresponding directions, when they are drawn from the given point P, we have

$$d\omega = d\psi \sin\theta\, d\theta\,. \tag{2.7}$$

For the problems discussed in Part I, however, it may be assumed that the physical state of the photospheric layers depends only on their depth, i. e. on the distance from the centre of the star. Owing to symmetry, the intensity of radiation in this case cannot depend on the azimuthal angle ψ. Consequently, the expression (2.7) for the solid angle can be immediately integrated over all angles ψ. We then obtain

$$d\omega = 2\pi \sin\theta\, d\theta\,. \tag{2.8}$$

The flux at a given point of the star, passing radially through unit area perpendicular to the radius, is then given by (2.5) as

$$\pi H_\nu = 2\pi \int_0^\pi I_\nu(\theta) \cos\theta \sin\theta\, d\theta\,. \tag{2.9}$$

The expression (2.6) is changed similarly.

In this particular case, which we shall usually be considering, the physical meaning of the flux is as follows. If the unit area is placed as shown in Fig. 2, i. e. with its plane perpendicular to the radius, then the flux is the excess of the energy passing upwards through this area ($0 \leq \theta \leq \frac{1}{2}\pi$) over the energy passing downwards ($\frac{1}{2}\pi \leq \theta \leq \pi$).

One of the chief differences between flux and intensity is that in "empty" space (i. e. when there is no absorption of radiation), the intensity of a ray remains constant along its path in space. For instance, at the outer boundary of the Earth's atmosphere radiation arrives from the Sun which is of the *same* intensity as that at the surface of the Sun itself (the absorption in interplanetary space being negligible). On the

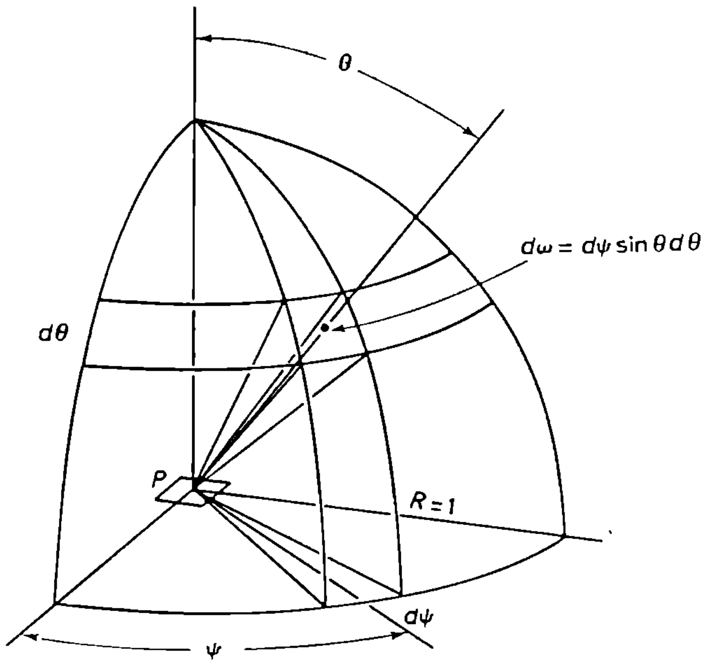

FIG. 2

other hand, the flux of solar (or stellar) radiation decreases with increasing distance from the Sun (or star), in inverse proportion to the square of the distance from the centre of the Sun (or star).

3. **The coefficients of emission and absorption.** We now consider the following question. Let an element of mass dm emit thermal energy in all directions. Then in time dt and frequency interval ν to $\nu + d\nu$ this element will emit within the solid angle $d\omega$ an amount of energy

$$j_\nu \, dm \, d\omega \, dt \, d\nu \, . \tag{2.10}$$

The coefficient of proportionality j_ν is called the **coefficient of emission.** It follows from the definition (2.10) that j_ν is the energy emitted in unit time, unit frequency interval and unit solid angle by unit mass of material.

The total energy emitted by the mass element dm in time dt is

$$dt \, dm \int_0^\infty \int_{4\pi} j_\nu \, d\omega \, d\nu \, . \tag{2.11}$$

In the case where the coefficient j_ν does not depend on direction, (2.11) may be rewritten

$$\mathrm{d}t\, \mathrm{d}m\, 4\pi \int_0^\infty j_\nu\, \mathrm{d}\nu\,. \tag{2.12}$$

We now introduce the **coefficient of absorption.** Let a beam of light rays whose intensity is I_ν fall perpendicularly on the surface of an absorbing layer of thickness $\mathrm{d}s$. As a result of passing through the layer, the intensity I_ν of the beam receives a (negative) increment $\mathrm{d}I_\nu$. The coefficient of absorption $\varkappa_\nu$ is defined as follows:

$$\mathrm{d}I_\nu = -\, I_\nu\, \varkappa_\nu\, \varrho\, \mathrm{d}s\,, \tag{2.13}$$

where ϱ is the density of the *absorbing** material. Thus the attenuation of the beam is, by (2.13), proportional to the intensity itself, the coefficient of proportionality being the coefficient of absorption. The correctness of the definition (2.13) is confirmed by experiment.

We now assume that the absorbing layer has a linear thickness s, and calculate the intensity of the beam emerging from the layer as a function of the other parameters. Dividing both sides of (2.13) by I_ν and integrating, we obtain

$$\log_e I_\nu = -\int_0^s \varkappa_\nu\, \varrho\, \mathrm{d}s + \log_e C\,, \tag{2.14}$$

where $\log_e C$ is a constant of integration. The solution (2.14) can be rewritten in the form

$$I_\nu = C \exp\left[-\int_0^s \varkappa_\nu\, \varrho\, \mathrm{d}s\right]. \tag{2.15}$$

The constant C is determined as follows. If s is put equal to zero, I_ν becomes the intensity of the *incident* beam, which we denote by I_ν^0. Therefore

$$I_\nu = I_\nu^0 \exp\left[-\int_0^s \varkappa_\nu\, \varrho\, \mathrm{d}s\right]. \tag{2.16}$$

Similar formulae are obtained if we consider not the intensity I_ν but the energy E_ν. Consequently, the incident beam is attenuated, in passing through the layer, according to an *exponential* law; the index of the exponent is

$$\tau_\nu = \int_0^s \varkappa_\nu\, \varrho\, \mathrm{d}s\,, \tag{2.17}$$

* The quantity ϱ can also be the total density of the material, in which case the quantity $\varkappa_\nu$ must be correspondingly changed; we shall discuss this later.

and this is called the **optical thickness** or **optical depth** of the layer. The optical thickness is a measure of the absorption of light in the given absorbing layer. According to (2.17), the element of optical thickness is

$$d\tau_\nu = \varkappa_\nu \varrho \, ds \, . \tag{2.18}$$

4. **The equation of transfer of radiation.** The equation (2.13) determines the change of I_ν over the path ds only on account of absorption. If the material not only absorbs but also emits energy, the equation (2.13) must be correspondingly modified. In this case we arrive at what is called the **equation of transfer of radiation.**

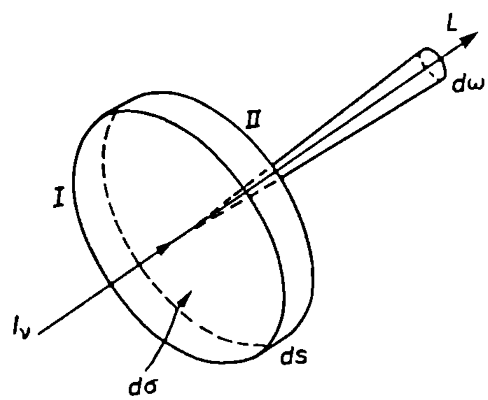

FIG. 3

We consider an elementary cylinder* with base area $d\sigma$ and height ds, whose axis is oriented along some direction L (Fig. 3). We denote the density of absorbing material in this cylinder by ϱ.

Let E_i be the quantity of radiant energy† passing in time dt through the first (I) base $d\sigma$ of the cylinder, inside the solid angle $d\omega$ and within the frequency interval ν to $\nu + d\nu$. This quantity, by (2.2), is ($\cos \theta = 1$)

$$E_i = I_\nu \, d\omega \, d\sigma \, dt \, d\nu \, . \tag{2.19}$$

The radiation emerging from the cylinder through its second (II) base, E_e, is characterised by a different intensity $I_\nu + dI_\nu$, and we have

$$E_e = (I_\nu + dI_\nu) \, d\omega \, d\sigma \, dt \, d\nu \, . \tag{2.20}$$

The reasons for the change of E_i to E_e are as follows. Firstly, the corresponding beam of energy emerging from the second base of the cylinder must be increased because *the material of the cylinder itself emits*, and in particular emits in the direction L. To take account of this, we must add to the quantity E_i the energy emitted by the cylinder in the direction L inside the same solid angle $d\omega$, in the same time interval dt and in the same frequency interval $d\nu$. We call this energy ΔE_{em}. By (2.10), it is equal to

$$\Delta E_{em} = j_\nu \varrho \, d\sigma \, ds \, d\omega \, dt \, d\nu \, , \tag{2.21}$$

since

$$dm = \varrho \, d\sigma \, ds \, . \tag{2.22}$$

---

* It is supposed that the height of the cylinder considered is negligible compared with the radius of its base.

† We omit the suffix ν to E for the sake of brevity.

The second reason for the change in E_i is the *absorption* of radiation in the cylinder. The quantity of energy absorbed, ΔE_{ab}, is, by (2.13),

$$\Delta E_{\mathrm{ab}} = -E_i\,\varkappa_\nu\,\varrho\,\mathrm{d}s\,, \tag{2.23}$$

since the definition (2.13) holds not only for intensity, but also for energy, as we have already remarked.

The equation which takes account of all changes in the quantity of energy which occur while the radiation passes through the cylinder is

$$E_e = E_i + \Delta E_{\mathrm{em}} + \Delta E_{\mathrm{ab}}\,, \tag{2.24}$$

where the negative sign of ΔE_{ab} has already been included in (2.23); this latter term may be rewritten, according to (2.19) and (2.23), as

$$\Delta E_{\mathrm{ab}} = -I_\nu\,\varkappa_\nu\,\varrho\,\mathrm{d}s\,\mathrm{d}\omega\,\mathrm{d}\sigma\,\mathrm{d}t\,\mathrm{d}\nu\,. \tag{2.25}$$

Applying the formulae (2.19), (2.20), (2.21) and (2.25), and effecting the necessary cancelling in the equation obtained. we find the required **equation of transfer:**

$$\mathrm{d}I_\nu/\mathrm{d}s = -I_\nu\,\varkappa_\nu\,\varrho + j_\nu\,\varrho\,. \tag{2.26}$$

This equation describes the changes in the intensity of radiation which occur when it passes through an absorbing and emitting medium.

The independent variable s in the equation (2.26) is in some ways inconvenient. We therefore transform to another variable, namely, the *depth* of the given layer in the stellar photosphere. In this connection we recall that we are investigating stellar photospheres whose thickness is very small compared with the radius of the star. Consequently, for every element of the stellar photosphere we can consider the photospheric layers as plane-parallel. The normal n to these layers (which coincides in direction with the radius of the star) and the direction L of the ray are at an angle θ (Fig. 4). We now introduce the depth h, reckoned along the normal and increasing inwards to the centre of the star. It then follows from Fig. 4 that

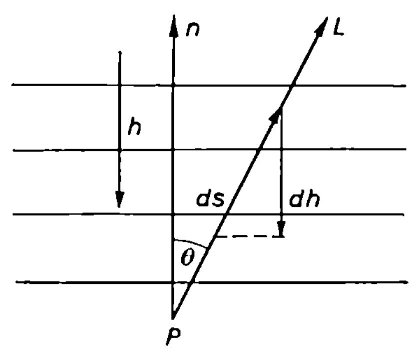

FIG. 4

$$\mathrm{d}s = -\sec\theta\,\mathrm{d}h\,. \tag{2.27}$$

From (2.27), the equation of transfer (2.26) can be rewritten as

$$\cos\theta\,\mathrm{d}I_\nu(\theta)/\mathrm{d}h = I_\nu(\theta)\,\varkappa_\nu\,\varrho - j_\nu\,\varrho\,, \tag{2.28}$$

and it is clear from (2.28) that I_ν is a function of the angle θ.

Chapter 3. Radiative equilibrium of the stellar photosphere. The solution of the equation of transfer

1. **The condition of radiative equilibrium.** We have already remarked that the principal problem of our theory is to establish the law of variation of temperature and other parameters (pressure, density, etc.) inside the stellar photosphere. Let us consider some element of volume Δv inside the stellar photosphere. Let the total quantity of energy *gained* per second by this element as a result of its interaction with other parts of the photosphere be E_+, and the total quantity of energy *lost* per second by the same element be E_-.

The temperature inside Δv will be determined by the relation between the quantities E_+ and E_-. If, for instance, $E_+ > E_-$, the internal energy of the volume will continually increase, and consequently the temperature inside it will increase. If, on the other hand, $E_+ < E_-$, then more energy will be lost than is gained, and consequently the temperature inside the element Δv will fall. We have stated in Chapter 1 that we intend to study the photospheres only of non-variable stars. In this case, the temperature inside any volume of the stellar photosphere must remain constant and independent of time. This will happen only if the following equality holds:

$$E_+ = E_- , \tag{3.1}$$

and this is the **condition of energy equilibrium.**

The equation (3.1) has a very general form. However, as we have stated in Chapter 2, the principal factor which determines the transfer of energy from point to point in the photospheres of non-variable stars is the transfer of thermal energy *by radiation*. Consequently, the quantities E_+ and E_- are, in the theory which we consider, identical with the energies of radiation: E_+ is the *radiant* energy *absorbed* per second by the volume Δv, and E_- is the *radiant* energy *emitted* per second by the same volume. In accordance with the above remarks, we shall henceforward call this theory the **theory of radiative equilibrium.**

Our first problem is to write down the quantities E_+ and E_- in an explicit form, and we take first E_-. Let the volume Δv considered have a mass $\mathrm{d}m$. Then the total quantity of energy emitted by the volume Δv in unit time is, by (2.10),

$$E_- = \mathrm{d}m \int_0^\infty \int_{4\pi} j_\nu \, \mathrm{d}\omega \, \mathrm{d}\nu , \tag{3.2}$$

where the integration is extended over the whole spectrum and over the whole solid angle 4π.

Next, to calculate E_+, we take the same element of volume Δv, with density of absorbing material ϱ. We surround this element (Fig. 5) by a closed surface Σ, whose dimensions are very large compared with those of the surface σ which bounds the volume Δv.

We also dispose the volume Δv in such a way that the distance from it to any point of the surface Σ is very large compared with the dimensions of the element Δv itself. We now consider a beam of rays entering the volume bounded by Σ and passing through the elements $d\Sigma$ and $d\sigma$ of the surfaces Σ and σ. Let the element $d\Sigma$ subtend a solid angle $d\omega$ at $d\sigma$ (and therefore at any point of the volume Δv, on account of the smallness of the latter). Then the quantity of radiant energy passing through the area $d\sigma$ in time dt, inside the solid angle $d\omega$ and the frequency interval from ν to $\nu + d\nu$ is, by (2.2),

$$I_\nu \, d\sigma \cos\theta \, dt \, d\omega \, d\nu \, , \tag{3.3}$$

and all this energy passes also through the area $d\Sigma$.

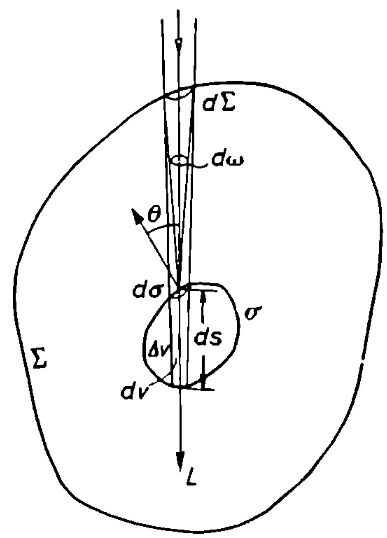

FIG. 5

Of the quantity of energy (3.3), the amount absorbed in the volume Δv is, by (2.13),

$$I_\nu \, d\sigma \cos\theta \, dt \, d\omega \, d\nu \, \varkappa_\nu \, \varrho \, ds \, , \tag{3.4}$$

where ds is the height of the cylinder* dv. The product $d\sigma \cos\theta \, ds$ is the volume of the cylinder dv, whose height is ds and base area $d\sigma \cos\theta$ (see Fig. 5). If we replace the volume $d\sigma \cos\theta \, ds$ by dv, the expression (3.4) becomes

$$I_\nu \, \varkappa_\nu \, \varrho \, dv \, dt \, d\omega \, d\nu \, . \tag{3.5}$$

We now integrate (3.5) over the whole volume Δv, and thus obtain the amount of radiant energy, absorbed in the whole volume Δv, which has first passed through $d\Sigma$. The quantity I_ν may be regarded as constant in the integration, because of the smallness of Δv and its remoteness from the surface Σ; we thus obtain simply Δv instead of dv in the expression (3.5).

To take account of the radiation which falls on Δv from the whole surface Σ, we must integrate over ω the expression which we have obtained. In so doing, the position in Δv of the point from which the

* The volume dv may be regarded as a cylinder in consequence of the smallness of $d\omega$.

solid angles $d\omega$ are measured is quite immaterial, since we again suppose Δv small and very remote from the surface Σ. We find for the total energy $(E_+)_\nu \, d\nu$ absorbed per second by the volume Δv and included in the frequency interval $d\nu$ (using $dm = \varrho \, \Delta v$)

$$(E_+)_\nu \, d\nu = \varkappa_\nu \, dm \, d\nu \int_{4\pi} I \, d\omega \,. \tag{3.6}$$

Finally, integrating over all frequencies, we obtain

$$E_+ = dm \int_0^\infty \int_{4\pi} I_\nu \, \varkappa_\nu \, d\omega \, d\nu \,. \tag{3.7}$$

Consequently, we may write the *condition of radiative equilibrium*, on the basis of (3.2) and (3.7), as

$$\int_0^\infty \int_{4\pi} j_\nu \, d\omega \, d\nu = \int_0^\infty \int_{4\pi} I_\nu \, \varkappa_\nu \, d\omega \, d\nu \,. \tag{3.8}$$

In the presence of radiative equilibrium the integrated flux of radiation πH in a "plane" photosphere remains *constant* at all depths. To show this, we integrate the equation of transfer (2.28) over the whole sphere and over the whole spectrum. We then find

$$\frac{d}{dh} \int_0^\infty \int_{4\pi} I_\nu \, (\theta) \cos \theta \, d\omega \, d\nu = \varrho \int_0^\infty \int_{4\pi} [I_\nu \, (\theta) \, \varkappa_\nu - j_\nu] \, d\omega \, d\nu \,, \tag{3.9}$$

since the variables ω, ν and h are independent. The quantity under the differentiation sign is, by (2.6), the flux πH. But the right-hand side of (3.9) is zero, by (3.8). Thus the flux πH is *constant* (independent of h).

The following two points should be noted in connection with this result:

(1) The constancy of the flux πH occurs only in photospheres whose layers may be regarded as plane-parallel. In fairly extended photospheres (which we do not consider at this stage), the flux πH decreases outwards in inverse proportion to the square of the distance from the centre of the star.

(2) The result that the total flux πH is constant for plane photospheres can by no means be extended to the monochromatic flux πH_ν, which in general varies with depth. Since, however, the flux πH is constant, this means that its spectral composition varies with depth. In particular, as we penetrate deeper into the photosphere, radiation of short wavelengths plays an increasing part. We shall give in Chapter 6 an actual example of the depth dependence of πH_ν for different frequencies.

2. The hypothesis of local thermodynamic equilibrium. We now introduce into the basic equations (2.28) and (3.8) the *temperature* — the principal physical parameter of the problem. We suppose (and this will be the fundamental hypothesis of our theory) that in any small region of the stellar photosphere the ratio of the coefficients j_ν and $\varkappa_\nu$ is the same as in thermodynamic equilibrium. In this case, by Kirchhoff's Law,

$$j_\nu = \varkappa_\nu\, B_\nu(T)\,, \tag{3.10}$$

where $B_\nu\,(T)$ is the intensity of radiation of a black body, determined by Planck's Law:

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\,\frac{1}{\exp\,(h\nu/kT) - 1}\,. \tag{3.11}$$

Thus we suppose that the law (3.10) may be applied to any small region of a stellar photosphere; we identify the parameter T which appears in this equation with the temperature at the given level in the photosphere. We call this supposition the hypothesis of **local thermodynamic equilibrium,** as distinct from the equilibrium case where the temperature is constant *throughout* the region considered.

To justify the hypothesis which we have introduced about the existence of local thermodynamic equilibrium in the photospheres of stars, we can proceed in two ways: firstly, by using (3.10) and completing our theory of radiative equilibrium, and then comparing its results with those of observation, which is a test of the correctness of the above hypothesis; secondly, by developing the theory of stellar photospheres and determining the physical conditions in the photospheres of stars, we can test the applicability of the law (3.10) from a purely physical point of view. In what follows we shall use both methods. It must be kept in mind, however, that the supposition that local thermodynamic equilibrium, for the appropriate temperature, exists in every small region of the stellar photosphere is as yet only a hypothesis, so far as we are concerned.

Introducing Kirchhoff's Law (3.10) into the equation of transfer (2.28) we find

$$\cos\theta\, \mathrm{d}I_\nu(\theta)/\mathrm{d}\tau_\nu = I_\nu(\theta) - B_\nu\,, \tag{3.12}$$

where

$$\mathrm{d}\tau_\nu = \varkappa_\nu\, \varrho\, \mathrm{d}h \tag{3.13}$$

is the element of optical thickness τ_ν. It is to be noticed that we have introduced the variable τ_ν in place of the depth h. This variable is completely determined by (3.13) if a definite zero point of τ_ν is specified. It is convenient to reckon τ_ν as zero at the *outer boundary* of the atmosphere of the star. Then $\tau_\nu = \int\limits_0^h \varkappa_\nu\, \varrho\, \mathrm{d}h$.

Next, introducing (3.10) into the condition of radiative equilibrium (3.8), we obtain

$$\int_0^\infty \int_{4\pi} I_\nu(\theta)\, \varkappa_\nu \,\mathrm{d}\omega\, \mathrm{d}\nu = \int_0^\infty \int_{4\pi} B_\nu\, \varkappa_\nu\, \mathrm{d}\omega\, \mathrm{d}\nu\, . \tag{3.14}$$

The condition of radiative equilibrium (3.14) may be further rewritten as follows. Introducing the new quantity

$$J_\nu = \int_{4\pi} I_\nu(\theta) \frac{\mathrm{d}\omega}{4\,\pi}\, , \tag{3.15}$$

which is the **mean intensity** of radiation at a given point, dividing (3.14) by $4\,\pi$, and noticing that $B_\nu(T)$ is independent of θ, we find

$$\int_0^\infty J_\nu\, \varkappa_\nu\, \mathrm{d}\nu = \int_0^\infty B_\nu\, \varkappa_\nu\, \mathrm{d}\nu\, . \tag{3.16}$$

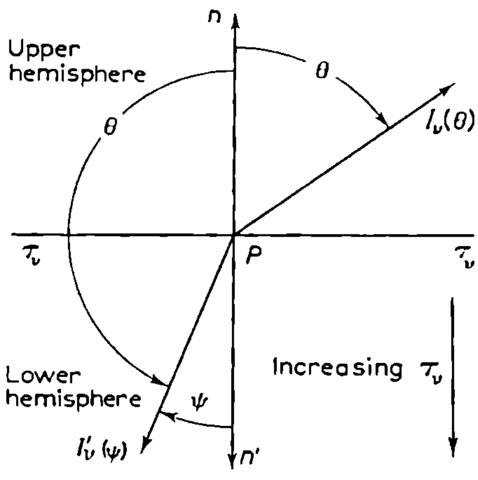

FIG. 6

3. The solution of the equations of transfer. We must now consider the solution of the equation of transfer (3.12). We divide the aggregate of directions of radiation at each point of the stellar photosphere into two groups: the upper and lower hemispheres (Fig. 6). The intensities corresponding to directions included in $0 \leq \theta \leq \frac{1}{2}\,\pi$ we shall call $I_\nu(\theta)$ as before. In the lower hemisphere $(\frac{1}{2}\,\pi \leq \theta \leq \pi)$, we introduce a new reckoning of angles, from the lower normal n', $\psi = \pi - \theta$ and denote the intensity by $I'_\nu(\psi)$. The equation of transfer (3.12) for the lower hemisphere becomes

so

$$\cos\psi\, \mathrm{d}I'_\nu(\psi)/\mathrm{d}\tau_\nu = - I'_\nu(\psi) + B_\nu\, , \tag{3.17}$$

since $\cos\theta = -\cos\psi$.

The equations (3.12) and (3.17) may also be rewritten

$$\mathrm{d}I_\nu(\theta)/\mathrm{d}\tau_\nu - I_\nu(\theta)\sec\theta + B_\nu \sec\theta = 0\, , \tag{3.18}$$

$$\mathrm{d}I'_\nu(\psi)/\mathrm{d}\tau_\nu + I'_\nu(\psi)\sec\psi - B_\nu \sec\psi = 0\, . \tag{3.19}$$

Both these equations are linear equations of the form

$$\mathrm{d}y/\mathrm{d}x + P(x)\, y + Q(x) = 0\, . \tag{3.20}$$

The solution of such an equation can be written in either of the following two forms, which are completely equivalent:

$$\left.\begin{aligned} y &= C \exp\left[-\int_{x_0}^{x} P(x)\,\mathrm{d}x\right] + \\ &\quad + \exp\left[-\int_{x_0}^{x} P(x)\,\mathrm{d}x\right] \int_{x}^{\infty} Q(x) \exp\left[\int_{x_0}^{x} P(x)\,\mathrm{d}x\right] \mathrm{d}x\,, \\ y &= D \exp\left[-\int_{x_0}^{x} P(x)\,\mathrm{d}x\right] - \\ &\quad - \exp\left[-\int_{x_0}^{x} P(x)\,\mathrm{d}x\right] \int_{x_0}^{x} Q(x) \exp\left[\int_{x_0}^{x} P(x)\,\mathrm{d}x\right] \mathrm{d}x\,, \end{aligned}\right\} \tag{3.21}$$

where x_0 is a fixed arbitrary number, and C (or D) is a constant of integration. It is easy to see by direct substitution that both forms of the solution (3.21) satisfy the equation (3.20).

From comparison of (3.20) and (3.18) we have

$$y = I_\nu(\theta)\,, \quad x = \tau_\nu\,, \quad P(x) = -\sec\theta\,, \quad Q(x) = B_\nu \sec\theta\,, \tag{3.22}$$

and from comparison of (3.20) and (3.19)

$$y = I'_\nu(\psi)\,, \quad x = \tau_\nu\,, \quad P(x) = \sec\psi\,, \quad Q(x) = -B_\nu \sec\psi\,. \tag{3.23}$$

In what follows, we shall take the zero point of optical depth at the outer boundary of the atmosphere of the star, i. e. in (3.21) the quantity x_0 must be put equal to zero. According to (3.13), τ_ν will increase towards the centre of the star (see Fig. 6). Consequently, we have for the solutions of equations (3.18) and (3.19)

$$\int_{x_0}^{x} P(x)\,\mathrm{d}x = -\int_{0}^{\tau_\nu} \sec\theta\,\mathrm{d}\tau_\nu = -\tau_\nu \sec\theta \tag{3.24}$$

and

$$\int_{x_0}^{x} P(x)\,\mathrm{d}x = \int_{0}^{\tau_\nu} \sec\psi\,\mathrm{d}\tau_\nu = \tau_\nu \sec\psi\,. \tag{3.25}$$

Using (3.21) — (3.25), we can write the solutions of equations (3.18) and (3.19) in the following form:

$$I_\nu(\theta,\tau_\nu) = C_\nu\, e^{\tau_\nu \sec\theta} + e^{\tau_\nu \sec\theta} \int_{\tau_\nu}^{\infty} B_\nu\, e^{-t_\nu \sec\theta} \sec\theta\,\mathrm{d}t_\nu\,, \tag{3.26}$$

$$I'_\nu(\psi,\tau_\nu) = D_\nu\, e^{-\tau_\nu \sec\psi} + e^{-\tau_\nu \sec\psi} \int_{0}^{\tau_\nu} B_\nu\, e^{t_\nu \sec\psi} \sec\psi\,\mathrm{d}t_\nu\,; \tag{3.27}$$

in writing (3.26) and (3.27) we have emphasised that the intensities are calculated for the level in the stellar photosphere to which the optical depth τ_ν corresponds. In order to distinguish this fixed optical depth from the variable which appears under the sign of integration, we have denoted the latter by t_ν.

We now consider the determination of the constants of integration C_ν and D_ν. The physical meaning of these constants will become clear if we analyse the meaning of each term in the solutions (3.26) and (3.27). The quantity of radiant energy, emitted in some direction θ in time dt, frequency interval (ν, $\nu + \mathrm{d}\nu$) and solid angle dω by a mass element ϱ ds (a cylinder) with unit base area (see Fig. 3) at a depth t_ν, is, by (2.10),

$$j_\nu \, \varrho \, \mathrm{d}s \, \mathrm{d}\omega \, \mathrm{d}t \, \mathrm{d}\nu \, . \tag{3.28}$$

On its way to a layer at a smaller depth τ_ν, this energy will be weakened by some exponential factor, according to (2.16). If the radiation travelled normally to the layers, the index of the exponential would be $-(t_\nu - \tau_\nu)$. But since the radiation travels at an angle θ to the normal, the index is

$$-(t_\nu - \tau_\nu) \sec \theta \, . \tag{3.29}$$

Hence the energy (3.28) emitted by the element will be reduced, at a depth τ_ν, to

$$j_\nu \, e^{-(t_\nu - \tau_\nu) \sec \theta} \, \varrho \, \mathrm{d}s \, \mathrm{d}\omega \, \mathrm{d}t \, \mathrm{d}\nu \, , \tag{3.30}$$

or, by (3.10), (3.13) and (2.27),

$$- B_\nu \, e^{-(t_\nu - \tau_\nu) \sec \theta} \sec \theta \, \mathrm{d}t_\nu \, \mathrm{d}\omega \, \mathrm{d}t \, \mathrm{d}\nu \, . \tag{3.31}$$

In order to take account of the energy arriving at the layer τ_ν in the given direction from all deeper layers, and not only from the depth t_ν, we must integrate the expression (3.31) over all t_ν from τ_ν to ∞. Consequently (3.31) becomes*

$$\mathrm{d}\omega \, \mathrm{d}t \, \mathrm{d}\nu \int_{\tau_\nu}^{\infty} B_\nu \, e^{-(t_\nu - \tau_\nu) \sec \theta} \sec \theta \, \mathrm{d}t_\nu \, . \tag{3.32}$$

On dividing this expression by dω dt dν, we find the intensity of radiation arriving from below at the layer of depth τ_ν in the direction θ:

$$I_\nu(\theta, \tau_\nu) = \int_{\tau_\nu}^{\infty} B_\nu \, e^{-(t_\nu - \tau_\nu) \sec \theta} \sec \theta \, \mathrm{d}t_\nu \, , \tag{3.33}$$

that is, the second term of (3.26).

* Since the quantity s increases outwards from the centre of the star, the integration of (3.31) must go from $t_\nu = \infty$ to $t_\nu = \tau_\nu$ (see Fig. 4). Interchanging the limits of integration and changing the sign, we obtain (3.32).

Thus the second term of (3.26) takes account of the fact that the radiation arriving at the layer of depth τ_ν, at an acute angle θ to the outward normal, is composed of radiation coming from all layers at depths $t_\nu > \tau_\nu$. The first term of (3.26) accounts for additional radiation arriving at the layer of depth τ_ν but not coming from layers at a finite depth. This term represents the existence at an infinite depth of a surface emitting radiation of infinite intensity, which decreases exponentially as we move outwards. In our physical problem there is no such surface, and therefore $C_\nu = 0$.

Similarly, it is easy to see that, for radiation travelling inwards at an acute angle ψ to the inward normal, the intensity at depth τ_ν is composed of the intensities from the various layers of the photosphere lying at depths $t_\nu < \tau_\nu$. This is the physical meaning of the second term of (3.27). The first term represents the radiation entering the photosphere from outside, which is attenuated by a factor $e^{-\tau_\nu \sec \psi}$ in reaching the depth τ_ν. In all cases, with the exception of close double stars, where the radiation of one star on to the photosphere of the other has a noticeable effect, the amount of radiation incident on the photosphere is negligible, and therefore we can put $D_\nu = 0$.

Thus we have

$$C_\nu = 0 \,, \tag{3.34}$$

$$D_\nu = 0 \,. \tag{3.35}$$

Substituting $C_\nu = 0$ in (3.26) and using (3.35), we obtain from (3.26) and (3.27)

$$I_\nu(\theta, \tau_\nu) = e^{\tau_\nu \sec \theta} \int_{\tau_\nu}^{\infty} B_\nu \, e^{-t_\nu \sec \theta} \sec \theta \, dt_\nu \,, \tag{3.36}$$

$$I'_\nu(\psi, \tau_\nu) = e^{-\tau_\nu \sec \psi} \int_{0}^{\tau_\nu} B_\nu \, e^{t_\nu \sec \psi} \sec \psi \, dt_\nu \,. \tag{3.37}$$

From formula (3.36) it follows, *inter alia*, that the intensity of radiation emerging from the surface of the photosphere at an angle θ is

$$I_\nu(\theta, 0) = \int_{0}^{\infty} B_\nu \, e^{-\tau_\nu \sec \theta} \sec \theta \, d\tau_\nu \,, \tag{3.38}$$

since $\tau_\nu = 0$ at the boundary of the photosphere.

4. The averaging of the equations of transfer over angles. Let us now consider the question of the averaging of the equations of transfer of radiation. From the mathematical point of view, both the theory of the radiative equilibrium of stellar photospheres and the theory of

absorption lines (which we shall discuss in Part II) are very complicated. One of the most serious difficulties is that, in solving the problem, it is necessary to take account of the change of intensity of radiation with direction, i. e. as a function of the angles θ and ψ at each point of the photosphere. If especially high accuracy is not required in the final results, it is possible to introduce some simplifying assumptions concerning the quantities $I_\nu(\theta)$and $I'_\nu(\psi)$, by averaging them over direction. Where possible, we shall estimate the error due to averaging by comparing the results obtained with those of a more rigorous treatment of the appropriate problems.

We repeat the equations (3.12) and (3.17):

$$0 \leq \theta \leq \tfrac{1}{2}\pi\,, \quad \cos\theta\, \mathrm{d}I_\nu(\theta)/\mathrm{d}\tau_\nu = I_\nu(\theta) - B_\nu\,, \tag{3.12}$$

$$0 \leq \psi \leq \tfrac{1}{2}\pi\,, \quad \cos\psi\, \mathrm{d}I'_\nu(\psi)/\mathrm{d}\tau_\nu = -I'_\nu(\psi) + B_\nu\,. \tag{3.17}$$

Multiplying both sides of these equations by $1/2\pi$ and integrating over the corresponding hemisphere, we obtain

$$0 \leq \theta \leq \tfrac{1}{2}\pi\,, \quad \frac{\mathrm{d}}{\mathrm{d}\tau_\nu} \int\limits_{2\pi} I_\nu(\theta) \cos\theta \frac{\mathrm{d}\omega}{2\pi} = \int\limits_{2\pi} I_\nu(\theta) \frac{\mathrm{d}\omega}{2\pi} - B_\nu\,, \tag{3.39}$$

$$0 \leq \psi \leq \tfrac{1}{2}\pi\,, \quad \frac{\mathrm{d}}{\mathrm{d}\tau_\nu} \int\limits_{2\pi} I'_\nu(\psi) \cos\psi \frac{\mathrm{d}\omega}{2\pi} = -\int\limits_{2\pi} I'_\nu(\psi) \frac{\mathrm{d}\omega}{2\pi} + B_\nu\,, \tag{3.40}$$

since B_ν is independent of θ and ψ. We now remove $\cos\theta$ and $\cos\psi$ from under the sign of integration in (3.39) and (3.40), using the theorem of the mean and denoting their average values by $\overline{\cos\theta}$ and $\overline{\cos\psi}$; it is clear that, since $I_\nu(\theta)$ and $I'_\nu(\psi)$ depend on τ_ν, $\overline{\cos\theta}$ and $\overline{\cos\psi}$ will in general depend on τ_ν also. However, assuming as an approximation that these two quantities are constant, we find

$$\overline{\cos\theta}\, \mathrm{d}I_\nu/\mathrm{d}\tau_\nu = I_\nu - B_\nu\,, \tag{3.41}$$

$$\overline{\cos\psi}\, \mathrm{d}I'_\nu/\mathrm{d}\tau_\nu = -I'_\nu + B_\nu\,, \tag{3.42}$$

where

$$I_\nu = \int\limits_{2\pi} I_\nu(\theta) \frac{\mathrm{d}\omega}{2\pi} \tag{3.43}$$

and

$$I'_\nu = \int\limits_{2\pi} I'_\nu(\psi) \frac{\mathrm{d}\omega}{2\pi} \tag{3.44}$$

are the *mean* intensities over the corresponding hemispheres.

In order to find the intensities I_ν and I'_ν from equations (3.41) and (3.42), we must determine $\overline{\cos\theta}$ and $\overline{\cos\psi}$. Neither of these quantities is a simple average of the cosine of the angle over the corresponding

hemisphere*. According to (3.39) and (3.40) the average values of cos θ and cos ψ are weighted means with weights $I_\nu(\theta)$ and $I'_\nu(\psi)$ respectively. Hence the determination of $\overline{\cos\theta}$ and $\overline{\cos\psi}$ must be performed taking account of the general properties of the radiation field in a stellar photosphere. For this purpose we abandon, for the moment, the division of the radiation field into two hemispheres and the averaging of the intensities, and reckon θ as varying from 0 to π. In this case the equation of transfer must be taken in the form (3.12). Multiplying the latter by cos $\theta/4\pi$ and integrating over the whole sphere, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\tau_\nu}\int\limits_{4\pi} I_\nu(\theta)\cos^2\theta\,\frac{\mathrm{d}\omega}{4\pi} = \int\limits_{4\pi} I_\nu(\theta)\cos\theta\,\frac{\mathrm{d}\omega}{4\pi} - \int\limits_{4\pi} B_\nu\cos\theta\,\frac{\mathrm{d}\omega}{4\pi}\,. \quad (3.45)$$

Since B_ν is independent of θ, it is easy to see from (2.8) that the second term on the right of (3.45) is zero. We next introduce the quantity

$$K_\nu = \int\limits_{4\pi} I_\nu(\theta)\cos^2\theta\,\frac{\mathrm{d}\omega}{4\pi}\,. \quad (3.46)$$

Taking into account the definition of the flux (2.5), we have from (3.45)

$$\frac{\mathrm{d}K_\nu}{\mathrm{d}\tau_\nu} = \frac{H_\nu}{4}\,. \quad (3.47)$$

This equation is exact, and does not involve any approximation. We can use it to determine the quantities $\overline{\cos\theta}$ and $\overline{\cos\psi}$, by expressing K_ν, H_ν and J_ν (introduced in (3.15)) in terms of mean intensities. Taking first K_ν, we find from (3.46) and (2.8)

$$\begin{aligned}
K_\nu &= \int\limits_{4\pi} I_\nu(\theta)\cos^2\theta\,\frac{\mathrm{d}\omega}{4\pi} = \tfrac{1}{2}\int\limits_0^\pi I_\nu(\theta)\cos^2\theta\sin\theta\,\mathrm{d}\theta \\
&= \tfrac{1}{2}\Big\{\int\limits_0^{\frac{1}{2}\pi} I_\nu(\theta)\cos^2\theta\sin\theta\,\mathrm{d}\theta + \int\limits_{\frac{1}{2}\pi}^{\pi} I_\nu(\theta)\cos^2\theta\sin\theta\,\mathrm{d}\theta\Big\} \\
&= \tfrac{1}{2}\Big\{\int\limits_0^{\frac{1}{2}\pi} I_\nu(\theta)\cos^2\theta\sin\theta\,\mathrm{d}\theta + \int\limits_0^{\frac{1}{2}\pi} I'_\nu(\psi)\cos^2\psi\sin\psi\,\mathrm{d}\psi\Big\} \\
&= \tfrac{1}{2}\Big\{\overline{I_\nu(\theta)}\int\limits_0^{\frac{1}{2}\pi}\cos^2\theta\sin\theta\,\mathrm{d}\theta + \overline{I'_\nu(\psi)}\int\limits_0^{\frac{1}{2}\pi}\cos^2\psi\sin\psi\,\mathrm{d}\psi\Big\}\,, \quad (3.48)
\end{aligned}$$

* The simple average of cos θ is $\int\limits_0^{\frac{1}{2}\pi}\cos\theta\,\frac{\mathrm{d}\omega}{2\pi} = \frac{1}{2}$, and $\overline{\cos\psi}$ is the same. Putting $\overline{\cos\theta} = \overline{\cos\psi} = \frac{1}{2}$, we obtain a very crude approximation known as Schwarzschild's approximation.

where $\overline{I_\nu(\theta)}$ and $\overline{I'_\nu(\psi)}$ are the mean values of the intensities in the corresponding hemispheres, according to the theorem of the mean.

If we assume that these mean values are equal to the corresponding means in equations (3.41) and (3.42), and calculate the integrals in the last braces in (3.48), each of which equals $\frac{1}{3}$, we find

$$K_\nu = \tfrac{1}{6}(I_\nu + I'_\nu)\,. \tag{3.49}$$

By an exactly similar method we can calculate H_ν and J_ν from (2.5) and (3.15):

$$H_\nu = I_\nu - I'_\nu\,, \tag{3.50}$$

$$J_\nu = \tfrac{1}{2}(I_\nu + I'_\nu)\,. \tag{3.51}$$

Thus our main simplifying assumption is that in calculating K_ν, H_ν and J_ν we assume that in all three cases the mean intensities $\overline{I_\nu(\theta)}$ and $\overline{I'_\nu(\psi)}$ are equal to the corresponding mean values I_ν and I'_ν in the equations (3.41) and (3.42).

We could have arrived at the same expressions (3.49) to (3.51) and the same equations (3.41) and (3.42) by assuming from the beginning that in their respective hemispheres the intensities $I_\nu(\theta)$ and $I'_\nu(\psi)$ are independent of direction. In this case (see Fig. 7), we have

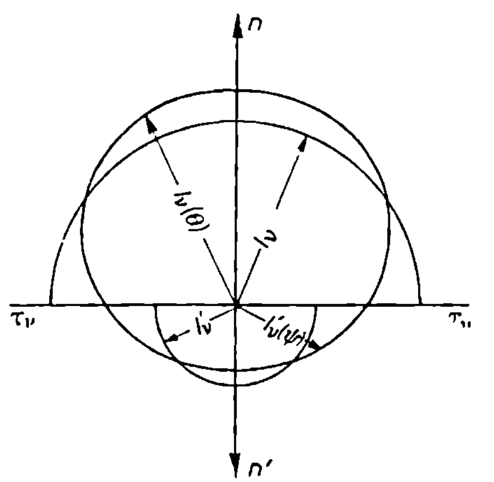

Fig. 7

$$0 \leq \theta \leq \tfrac{1}{2}\pi\,, \quad I_\nu(\theta) \equiv I_\nu\,, \tag{3.52}$$

$$0 \leq \psi \leq \tfrac{1}{2}\pi\,, \quad I'_\nu(\psi) \equiv I'_\nu\,. \tag{3.53}$$

Comparing (3.49) and (3.51), we find that

$$J_\nu = 3K_\nu\,, \tag{3.54}$$

and from this the equation (3.47) takes the form

$$\mathrm{d}J_\nu/\mathrm{d}\tau_\nu = \tfrac{3}{4}H_\nu\,. \tag{3.55}$$

From (3.50) and (3.51), this equation may also be written

$$\mathrm{d}(I_\nu + I'_\nu)/\mathrm{d}\tau_\nu = \tfrac{3}{2}(I_\nu - I'_\nu)\,. \tag{3.56}$$

Making use of the fact that the properties of the radiation field in the upper hemisphere cannot differ essentially from those in the lower hemisphere, we assume that in equations (3.41) and (3.42) the means of the cosines are equal to each other. This is our last simplification:

$$\overline{\cos\theta} = \overline{\cos\psi} = a\,. \tag{3.57}$$

Consequently equations (3.41) and (3.42) take the form

$$a\,\mathrm{d}I_\nu/\mathrm{d}\tau_\nu = I_\nu - B_\nu\,, \tag{3.58}$$

$$a\,\mathrm{d}I'_\nu/\mathrm{d}\tau_\nu = -I'_\nu + B_\nu\,. \tag{3.59}$$

Adding these, we have

$$a\,\mathrm{d}(I_\nu + I'_\nu)/\mathrm{d}\tau_\nu = I_\nu - I'_\nu\,. \tag{3.60}$$

Finally, comparing (3.60) and (3.56), we find that $a = \frac{2}{3}$, and thus our equations of transfer for the *mean* intensities take the final form

$$\tfrac{2}{3}\,\mathrm{d}I_\nu/\mathrm{d}\tau_\nu = I_\nu - B_\nu\,, \tag{3.61}$$

$$\tfrac{2}{3}\,\mathrm{d}I'_\nu/\mathrm{d}\tau_\nu = -I'_\nu + B_\nu\,. \tag{3.62}$$

The solution of these equations can be written down immediately. In fact, from a formal point of view, the last two equations can be considered as equations (3.12) and (3.17) with $\cos\theta = \cos\psi = \frac{2}{3}$. Since the solution of equations (3.12) and (3.17) is given by formulae (3.36) and (3.37), the solution of equations (3.61) and (3.62) is

$$I_\nu = e^{\frac{3}{2}\tau_\nu} \int\limits_{\tau_\nu}^{\infty} B_\nu\, e^{-\frac{3}{2}t_\nu}\, \tfrac{3}{2}\,\mathrm{d}t_\nu\,, \tag{3.63}$$

$$I'_\nu = e^{-\frac{3}{2}\tau_\nu} \int\limits_{0}^{\tau_\nu} B_\nu\, e^{\frac{3}{2}t_\nu}\, \tfrac{3}{2}\,\mathrm{d}t_\nu\,. \tag{3.64}$$

Chapter 4. The theory of radiative equilibrium for an absorption coefficient independent of the frequency

1. The variation of the temperature with depth in the photosphere. In the present chapter we shall consider the case where the absorption coefficient at every depth within the stellar photosphere is independent of the frequency. Matter which has this property is called **grey material.** It must be remarked at once that the case in question is very often extremely far from reality. Even in stars of the Sun's type, where the deviation from "greyness" is least, the absorption coefficient, as we shall see later, varies noticeably with frequency. Nevertheless, the study of the properties of a photosphere with $\varkappa_\nu = \varkappa =$ constant is of great interest in many respects. Firstly, in this case we can most easily estimate the error involved in the averaging of the intensities $I_\nu(\theta)$ and $I'_\nu(\psi)$ and their replacement by the quantities I_ν and I'_ν. Secondly, the case where $\varkappa_\nu = \varkappa =$ constant is in itself a standard, comparison with which makes it possible to estimate the dependence of $\varkappa_\nu$ on the frequency. Finally, the case $\varkappa_\nu = \varkappa$ was historically the first to be considered, and with it began the development of the theory of radiative equilibrium.

Let us consider first the question of the temperature distribution in the photosphere. Since the absorption coefficient is independent of the frequency, i. e. the condition

$$\varkappa_\nu = \varkappa = \text{constant} \tag{4.1}$$

holds, we can write for the optical depth

$$\mathrm{d}\tau = \varkappa \varrho \,\mathrm{d}h\,. \tag{4.2}$$

The condition of radiative equilibrium (3.16) takes, by virtue of (4.1), the form

$$J = B\,, \tag{4.3}$$

where

$$J = \int_0^\infty J_\nu \,\mathrm{d}\nu \tag{4.4}$$

and

$$B = \int_0^\infty B_\nu \,\mathrm{d}\nu\,. \tag{4.5}$$

Replacing the differential $\mathrm{d}\tau_\nu$ by $\mathrm{d}\tau$ in the equations of transfer (3.61) and (3.62) and integrating both equations over the frequency from $\nu = 0$ to $\nu = \infty$, we obtain

$$\tfrac{2}{3}\,\mathrm{d}I/\mathrm{d}\tau = I - B\,, \tag{4.6}$$

$$\tfrac{2}{3}\,\mathrm{d}I'/\mathrm{d}\tau = -I' + B\,, \tag{4.7}$$

where

$$I = \int_0^\infty I_\nu \,\mathrm{d}\nu\,, \quad \text{and } I' = \int_0^\infty I'_\nu \,\mathrm{d}\nu\,. \tag{4.8}$$

Integrating the equations (3.50) and (3.51) over the frequency between the same limits, we find

$$H = I - I'\,, \tag{4.9}$$

$$J = \tfrac{1}{2}(I + I')\,. \tag{4.10}$$

We now add (4.6) and (4.7), taking account of (4.9) and (4.10). We then obtain

$$\mathrm{d}J/\mathrm{d}\tau = \tfrac{3}{4}\,H\,. \tag{4.11}$$

Next, replacing the quantity J by B in accordance with (4.3) and taking account of the constancy of the flux H with depth, which we proved in Chapter 3, we can integrate equation (4.11), and this gives

$$B = \tfrac{3}{4}\,H\tau + B_0\,, \tag{4.12}$$

where B_0 is a constant of integration, which we can determine as follows. Since at the boundary of the star, where $\tau = 0$, the incident radiation vanishes, we have from (4.9)

$$H = I_0 \,. \tag{4.13}$$

Next, for $\tau = 0$ we have according to (3.63)

$$(I_\nu)_0 = \int_0^\infty B_\nu \, e^{-\frac{3}{2} t} \, \tfrac{3}{2} \, \mathrm{d}t \,, \tag{4.14}$$

or, integrating over the whole spectrum,

$$I_0 = \int_0^\infty B \, e^{-\frac{3}{2} t} \, \tfrac{3}{2} \, \mathrm{d}t \,. \tag{4.15}$$

Identifying (4.13) and (4.15) and introducing in the integral the solution (4.12), we have

$$H = \int_0^\infty (\tfrac{3}{4} H t + B_0) \, e^{-\frac{3}{2} t} \, \tfrac{3}{2} \, \mathrm{d}t \,, \tag{4.16}$$

where we have taken account of the fact that the function B in the integral depends on the variable of integration t. Next, replacing the quantity $\frac{3}{2}t$ in (4.16) by x, and using the general relation

$$\int_0^\infty e^{-x} \, \mathrm{d}x = \int_0^\infty x \, e^{-x} \, \mathrm{d}x = \int_0^\infty \frac{x^2 e^{-x}}{2!} \, \mathrm{d}x = \ldots = 1 \,, \tag{4.17}$$

we find at once

$$B_0 = \tfrac{1}{2} H \,. \tag{4.18}$$

Consequently, according to (4.12) and (4.18),

$$B = \tfrac{1}{2} H (1 + \tfrac{3}{2} \tau) \,. \tag{4.19}$$

We must now transform this to give the temperature. According to the theory of thermal radiation,

$$\pi B = \pi \int_0^\infty B_\nu \, \mathrm{d}\nu = \sigma T^4 \,, \tag{4.20}$$

where σ is a constant equal to $5{\cdot}672 \times 10^{-5}$ erg cm^{-2} sec^{-1} degree^{-4}. We now introduce what is called the **effective temperature** T_e of a star. It is defined in terms of the flux, as follows:

$$\pi H = \sigma T_e^4 \,. \tag{4.21}$$

Thus the product $\sigma\, T_e^4$ gives the number of ergs emitted in one second in all directions and over the whole spectrum by one square centimetre of the surface of the star. The introduction of the effective temperature by means of the relation (4.21) has a somewhat formal character. It is related to our previous remarks on the fact that stars radiate from their surfaces in accordance with Planck's Law (3.11). The effective temperature defines, in some sense, the *mean temperature* of the stellar photosphere and, by means of the relation (4.21), the *total flux* of radiation*.

By virtue of (4.21) the **absolute bolometric luminosity** of a star of radius R is

$$L = 4\,\pi\, R^2 \times \pi\, H = 4\,\pi\, R^2 \times \sigma\, T_e^4\,. \tag{4.22}$$

Introducing now (4.20) and (4.21) into (4.19), we find

$$T^4 = \tfrac{1}{2}\, T_e^4 (1 + \tfrac{3}{2}\,\tau)\,. \tag{4.23}$$

It follows from (4.23) that the boundary temperature T_0 of the photosphere, corresponding to $\tau = 0$, is determined by the relation

$$T_0^4 = \tfrac{1}{2}\, T_e^4\,. \tag{4.24}$$

From this it follows that $T_0 = 0{\cdot}841\ T_e$.

Moreover, we find from (4.23) that at the level $\tau = \frac{2}{3}$ the temperature of the photosphere is equal to the effective temperature.

Before drawing further conclusions based on the application of (4.19) or (4.23), we shall estimate the accuracy of these formulae. The theory of radiative equilibrium for the case of grey material is relatively simple, and thus it has been possible in this case to make considerable progress mathematically, without the introduction of auxiliary simplifying assumptions.

The mathematical analysis shows that, in the case considered, the exact relation between T_0 and T_e is

$$T_0^4 = \frac{\sqrt{3}}{4}\, T_e^4\,, \tag{4.25}$$

from which it follows that $T_0 = 0{\cdot}811\ T_e$. Moreover, S. Chandrasekhar, using the method of successive approximations developed by himself [29], has found in the fourth approximation, which is practically exact, the following dependence of B on τ:

$$B = \tfrac{3}{4}\, H (\tau + 0{\cdot}70692 - 0{\cdot}08392\, e^{-4{\cdot}45808\tau} - \\ - 0{\cdot}03619\, e^{-1{\cdot}50178\tau} - 0{\cdot}00946\, e^{-1{\cdot}10319\tau})\,, \tag{4.26}$$

and the relation (4.25) is satisfied here also.

* Of course, its spectral distribution may deviate quite considerably from the Planck distribution (3.11) with temperature $T = T_e$.

In order to compare the approximate solution (4.23) with the solution (4.26), we give (Fig. 8) the increase of T with τ, calculated on the basis of formulae (4.23) and (4.26) for a star with $T_e = 10{,}500°$. A comparison of the two curves shows that for many practical purposes we can restrict ourselves to the approximate solution (4.23). On the other hand, this approximation really depends only on the averaging of the intensities over direction. Consequently, it may be supposed that this process of averaging the intensities is suitable for many practical purposes in the more general case also, where the absorption coefficient depends on the frequency. This supposition is confirmed by the corresponding numerical calculations.

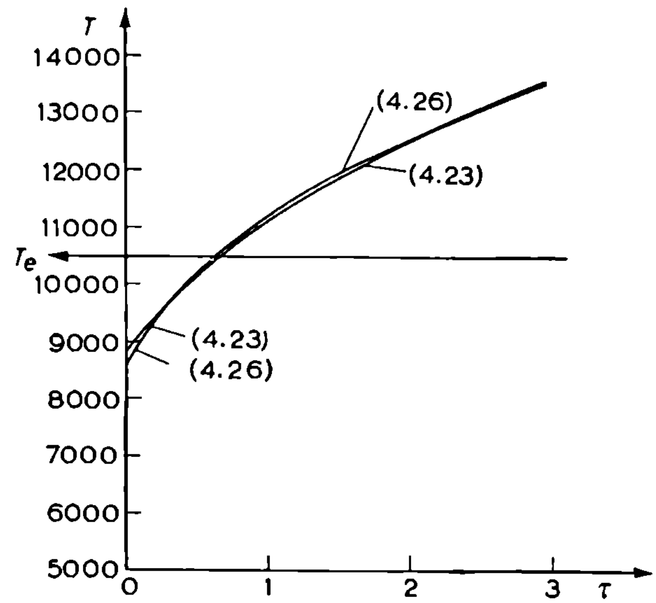

FIG. 8

2. **The law of darkening of the star's disc towards the limb.** Let us now consider the application of the formulae (4.19) and (4.23) which we have derived. We shall examine the dependence on direction of the radiation emerging from a star. Even a direct observation of the Sun's disc or of a photograph of it shows a decrease of the brightness of the Sun's light from the centre towards the limb. This is easily explained qualitatively.

We first write down the general expression for the intensity of radiation emerging from the surface of the Sun (or star) at an angle θ. According to (3.38), this intensity is, by (4.1) and (4.2),

$$I_\nu(\theta, 0) = \int_0^\infty B_\nu\, e^{-t \sec\theta} \sec\theta \, dt\,, \tag{4.27}$$

since at the boundary of the star $\tau_\nu = \tau = 0$.

The expression (4.27) can also be rewritten

$$I_\nu(\theta, 0) = \int_0^\infty B_\nu\, e^{-t \sec\theta} \sec\theta \, dt \Big/ \int_0^\infty e^{-t \sec\theta} \sec\theta \, dt\,, \tag{4.28}$$

since the denominator in (4.28) is equal to unity.

Thus the intensity $I_\nu(\theta,0)$ may be regarded as a weighted mean of the quantity B_ν. It also follows from (4.28) that, the larger sec θ, the smaller the optical depths from which radiation reaches us. In fact, even for small t considerably less than unity, the product t sec θ for $\theta \to \frac{1}{2}\pi$ becomes so large that the radiation from these depths does not emerge from the Sun at all (owing to the effect of the factor $e^{-t \sec \theta}$). And since B_ν increases with t, the radiation coming from the centre of the disc ($\theta = 0$) reaches the surface from hotter (i. e. deeper) layers than that coming from the outer parts of the disc. For this reason the inequality

$$I_\nu(\theta, 0) < I_\nu(0, 0) \tag{4.29}$$

must always hold.

In particular, it follows from (4.28) that, at the limb of the Sun's disc, where $\theta \to \frac{1}{2}\pi$ and sec $\theta \to \infty$, we have (independently of the absorption coefficient)

$$I_\nu(\tfrac{1}{2}\pi, 0) \approx B_\nu(T_0)\,. \tag{4.30}$$

Let us now consider this question quantitatively, taking first the total radiation. To do this, we integrate (4.27) over ν from 0 to ∞. We find

$$I(\theta, 0) = \int_0^\infty B\, e^{-t \sec \theta} \sec \theta \, dt\,. \tag{4.31}$$

Substituting (4.19) in (4.31) and effecting the integration involved (replacing c sec θ by x), we obtain

$$I(\theta, 0) = \tfrac{1}{2} H (1 + \tfrac{3}{2} \cos \theta)\,. \tag{4.32}$$

At the centre of the disc, where $\theta = 0$,

$$I(0, 0) = \tfrac{5}{2} \cdot \tfrac{1}{2} H\,, \tag{4.33}$$

whence

$$I(\theta, 0) = I(0, 0)\,(1 - \tfrac{3}{5} + \tfrac{3}{5} \cos \theta) = I(0, 0)\,(1 - u + u \cos \theta)\,, \tag{4.34}$$

where $u = 3/5 = 0.60$.

Observations, on the whole, confirm the law (4.34) and give $u \approx 0{\cdot}56$. Let us now consider the law of darkening towards the limb for radiation of wavelength λ. Since observers commonly use a wavelength scale, we rewrite (4.27) as follows:

$$I_\lambda(\theta,0) = \int_0^\infty B_\lambda\, e^{-t \sec \theta} \sec \theta \, dt\,. \tag{4.35}$$

The transformation from B_ν given by formula (3.11) to B_λ is effected as follows. Since the quantity of energy emitted by any source in a given spectrum interval is independent of the scale used, we may write

$$B_\nu \, | \, d\nu \, | = B_\lambda \, | \, d\lambda \, | \, , \tag{4.36}$$

and since

$$\nu = c/\lambda \, , \tag{4.37}$$

we have

$$| \, d\nu \, | = c \, | \, d\lambda \, | \, / \, \lambda^2 \, . \tag{4.38}$$

Using (3.11), (4.36), (4.37) and (4.38), we obtain

$$B_\lambda = \frac{2hc^2}{\lambda^5} \frac{1}{\exp(hc/k\lambda T) - 1} \, . \tag{4.39}$$

Introducing now (4.39) into (4.35) and using (4.23), we have

$$I_\lambda(\theta, 0) = \frac{2hc^2}{\lambda^5} \int\limits_0^\infty \frac{e^{-t \sec \theta} \sec \theta \, dt}{\exp[hc/k\lambda T_e(\frac{1}{2} + \frac{3}{4}t)^{\frac{1}{4}}] - 1} \, . \tag{4.40}$$

The integral in (4.40) may be calculated by numerical integration. Such calculations lead to the following results:

(1) The darkening to the limb increases with decreasing wavelength. In particular,

$$\lim_{\lambda \to 0} \frac{I_\lambda(\frac{1}{2}\pi, 0)}{I_\lambda(0, 0)} = 0 \, , \qquad \lim_{\lambda \to \infty} \frac{I_\lambda(\frac{1}{2}\pi, 0)}{I_\lambda(0, 0)} = 0{\cdot}817 \, , \tag{4.41}$$

where $I_\lambda(0,0)$ is the intensity of radiation coming from the centre of the Sun's disc, and $I_\lambda(\frac{1}{2}\pi,0)$ is the intensity of radiation coming from the extreme limb.

(2) The agreement of the above theory with observation is on the whole satisfactory; the greatest deviation does not exceed 10 to 15 %, and on the average it is considerably less.

3. The continuous spectrum of the star. In order to make sure of the correctness of our original assumptions, the agreement between the theory and observations of the darkening law is not in itself sufficient. We must also compare the theoretical and observed distribution of energy in the continuous spectrum of the Sun and of the stars. This we can do in two ways. For the Sun, we can study the distribution of energy in the continuous spectrum of any point of its disc. The theoretical distribution of energy in the continuous spectrum will be determined in this case by the expression (4.40). In fact, having fixed some value of θ in (4.40), we can study the variation of $I_\lambda(\theta,0)$ with wavelength for this θ. For the stars, however, this method is inapplicable, since

we cannot distinguish their discs, and only the total radiation sent out by the entire disc of the star in question is measured. Let us now calculate this radiation, assuming that the observer is located sufficiently far from the star.

The distance of any point on the disc of the star from its centre is denoted by r. Then the area of a concentric annulus of width $\mathrm{d}r$ and radius r is

$$\mathrm{d}S = 2\,\pi\, r\, \mathrm{d}r\ . \tag{4.42}$$

Next, let the radiation emerging from this annulus in the direction of the observer have an intensity $I_\nu(r, 0)$ (Fig. 9). Then this annulus sends to the observer in time $\mathrm{d}t$ the following quantity of energy:

$$\mathrm{d}E_\nu = 2\,\pi\, r\, I_\nu(r, 0)\, \mathrm{d}r\, \mathrm{d}\nu\, \mathrm{d}t\, \Delta\Omega\ , \tag{4.43}$$

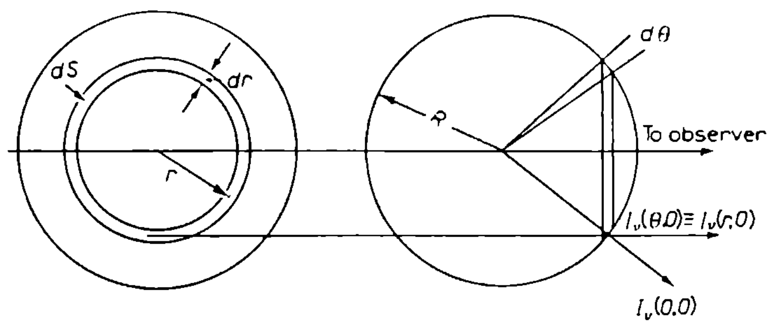

Fig. 9

where $\Delta\Omega$ is the solid angle subtended at the star by the observer's receiving apparatus. According to (4.43), the quantity of energy sent to the observer by the whole disc of the star is

$$\Delta E_\nu = \mathrm{d}\nu\, \mathrm{d}t\, \Delta\Omega\, 2\pi \int_0^R I_\nu(r, 0)\, r\, \mathrm{d}r\ , \tag{4.44}$$

where R is the radius of the star. Taking account also of the obvious relations

$$r = R \sin\theta\ , \quad \mathrm{d}r = R\cos\theta\ \mathrm{d}\theta\ , \tag{4.45}$$

we obtain

$$\begin{aligned} \Delta E_\nu &= \mathrm{d}\nu\, \mathrm{d}t\, \Delta\Omega\, 2\pi\, R^2 \int_0^{\frac{1}{2}\pi} I_\nu(\theta, 0) \sin\theta\cos\theta\, \mathrm{d}\theta \\ &= R^2\, \pi (H_\nu)_0\, \mathrm{d}\nu\, \mathrm{d}t\, \Delta\Omega\ . \end{aligned} \tag{4.46}$$

Thus we receive from the distant star radiation whose spectral composition is determined by the flux at its surface; the magnitude of this flux is denoted by $\pi (H_\nu)_0$.

If we now divide the expression (4.46) by πR^2 (the area of the star's disc) and by the product $d\nu\, dt\, \Delta\Omega$, we shall evidently obtain the mean intensity $\bar{I}_\nu$ of the radiation coming from the distant star. According to (4.46) we shall then have

$$\bar{I}_\nu = (H_\nu)_0 \,, \tag{4.47}$$

i. e. the mean intensity $\bar{I}_\nu$ equals the flux $(\pi H_\nu)_0$ divided by π.

To calculate $(\pi H_\nu)_0$ we can again use the approximate expressions of the last chapter. According to (3.50),

$$(H_\nu)_0 = (I_\nu)_0 \,, \tag{4.48}$$

since at the boundary $I'_\nu \equiv 0$. Consequently, introducing (3.63), we obtain

$$\bar{I}_\nu = (H_\nu)_0 = \int\limits_0^\infty B_\nu\, e^{-\frac{3}{2} t_\nu}\, \tfrac{3}{2}\, dt_\nu \,, \tag{4.49}$$

or, in the case of *grey material*,

$$\bar{I}_\nu = (H_\nu)_0 = \int\limits_0^\infty B_\nu\, e^{-\frac{3}{2} t}\, \tfrac{3}{2}\, dt \;. \tag{4.50}$$

The expressions (4.49) and (4.50) give the theoretical energy distribution in the continuous spectrum of distant stars.

Transforming now to the wavelength scale, and using (4.50), (4.39) and (4.23), we find

$$\bar{I}_\lambda = \frac{2hc^2}{\lambda^5} \int\limits_0^\infty \frac{e^{-\frac{3}{2} t}\, \frac{3}{2}\, dt}{\exp\left[hc/k\lambda\, T_e(\frac{1}{2} + \frac{3}{4} t)^{\frac{1}{4}}\right] - 1} \,. \tag{4.51}$$

On fixing a definite value for T_e, corresponding to the star in question, we can, as in the case of (4.40), find the dependence of the stellar (or solar) continuous spectrum radiation on wavelength. The appropriate calculations show that the theoretical energy distribution in the continuous spectrum of a star having a given effective temperature T_e is very close to the Planck distribution (4.39) for the *same* effective temperature T_e. The most important deviations of the law (4.51) from the Planckian one (4.39) for the same T_e are the following:

(1) The curve obtained from (4.51) is slightly displaced in the direction of *short* wavelengths, compared with the pure Planck curve (4.39).

(2) In the far ultra-violet region of the spectrum the Planckian character of the law (4.51) disappears, and $\bar{I}_\lambda$ is considerably greater than B_λ

calculated from formula (4.39) for the same λ and T_e. This latter deviation progressively increases with diminishing λ.

Generally, in the observed part of the spectrum of any star, the theoretical energy distribution given by (4.51) must approximately coincide with the Planck distribution (4.39), if the parameter T is taken equal to the *effective* temperature T_e of the star. This result also must be compared with observational data, and we shall be interested only in the relative distribution of energy, i. e. the dependence of the intensity of the stellar (or solar) radiation on the wavelength.

In order to be able to describe quantitatively the energy distribution in some wavelength interval, we introduce what is called the colour temperature. We find the value of the temperature $T = T_c$ for which the relative intensity distribution calculated from Planck's formula (4.39) best represents the observed intensity distribution in this interval in the star's spectrum. This temperature is called the **colour temperature** of the star [97]. In the continuous spectrum of a black body of temperature T, we have in any part of the spectrum $T_c = T$. On the other hand, for non-black radiation the parameter T_c may vary considerably with the wavelength interval for the same source of radiation.

Let us now turn to the observations, and consider first the radiation sent out by the whole disc of the Sun. According to existing observational data, the continuous spectrum of this radiation in the interval 4500 to 7000 Å can be approximately represented by a Planck curve with $T_c = 7000°$ to $7200°$, while the effective temperature T_e of the Sun is 5710°. There is an even greater divergence between T_e and T_c for stars of non-solar type.

The spectra of cool stars of classes M to N are so distorted by molecular absorption bands that the introduction of the colour temperature in these cases often proves to be quite meaningless.

Moreover, another phenomenon is observed in the interval of spectral classes from G 0 to B 0. Here the intensity of the continuous spectrum has a discontinuity at the limit of the Balmer series ($\lambda = 3646$ Å). Thus, for instance, in stars of class A 0 the intensity of radiation at this wavelength suddenly diminishes by a factor of three as we pass towards shorter wavelengths. The quantity T_c up to the limit of the Balmer series is different in value from that beyond. In particular, for these stars of class A 0 the value of T_c beyond the series limit is approximately 11,000°, while up to the series limit, i. e. from $\lambda = 3646$ Å towards longer wavelengths, the colour temperature in general varies with wavelength, being on the average 14,000° to 17,000°; the effective temperature of an A 0 class star, however, is close to 10,500°.

Thus there is a difference between T_c and T_e for all stars. Hence it follows that the theory explained in the present chapter must be

modified in some way, especially since the presence of a discontinuity at the limit of the Balmer series in the spectra of early-type stars is quite unintelligible from the standpoint of this theory (see Chapter 6).

Further development of the physics of stellar atmospheres has shown that the chief cause of these discrepancies is the incorrectness of the assumption that the absorption coefficient is independent of the frequency. In fact, the analysis of the physical conditions existing in stellar atmospheres shows that the absorption coefficient must vary with the frequency (or wavelength). Where this variation is least, the agreement between the theory with $\varkappa_\nu = \varkappa =$ constant and the observations should be best. In particular, the agreement between the theory considered and the observations regarding the darkening to the Sun's limb indicates that in the Sun's photosphere the deviations from $\varkappa_\nu = \varkappa =$ constant are not large.

In connection with what has been said above, we must consider the theory of radiative equilibrium for an absorption coefficient depending on the frequency. However, before doing so, we must acquaint ourselves with some subsidiary topics relating to the absorption coefficient.

Chapter 5. The coefficients of continuous absorption

1. **The chemical composition of stellar atmospheres.** The question of the chemical composition of stellar atmospheres is important, for the following reasons. Physical theory, confirmed by experiments, shows that the absorbing power of any material depends on the physical conditions in which the material is placed. It also appears that the absorbing power and its dependence on the physical conditions are different for different chemical elements. Hence the resulting absorption coefficient $\varkappa_\nu$, determined by the chemical composition of the whole photosphere, will depend markedly on the relative content of various elements in the photosphere of the star.

The chemical composition of stellar atmospheres is determined from a study of spectral lines; this topic will be considered in detail in Part II. Here we shall give some of the results, prefacing them with some necessary remarks.

The accumulated relevant astrophysical data indicate that the *chemical composition of the atmosphere* is approximately *the same* for the great majority of stars. Starting from this fact, it is possible to give a fully satisfactory explanation of the variations in the spectra of stars along the spectral sequence, assuming a change only of the physical conditions in passing from one star to another; agreement

between theory and observation can, indeed, be attained not only qualitatively, but also quantitatively. The identity of the chemical composition of the atmospheres of various main-sequence stars has recently been confirmed by a study of spectral lines (absorption lines). It is true that in some cases the chemical composition differs from the (so to speak) "standard" chemical composition. Deviations from this composition exist, apparently, in the atmospheres of cool stars where the bifurcation of the spectra into the M-branch and the R-N-branch occurs. Moreover, these deviations are considerable for Wolf-Rayet stars (carbon and nitrogen series) and for a number of stars with peculiar kinds of spectrum. However, we are not considering these stars at present, and we shall take as a starting value for the chemical composition the standard composition, using for this the results obtained for the Sun, whose chemical composition has been the most carefully and completely studied.

Since the density of material in the atmospheres of different stars is in general different, we shall have to speak of the *relative* chemical composition, i. e. of the percentage content of different elements in the stellar atmosphere. The relative content can be defined in two ways: (1) by taking the content of some element as unity and the content of the remaining elements relative to this element, (2) by taking the content of all elements as 100 % (or as unity) and giving the content of each element in per cent (or in fractions of unity).

The relative content may be given either *by number of atoms* or *by mass*. Let there be in 1 cm^3 of the stellar atmosphere n_1 atoms of the first element, n_2 of the second element, etc. Then the relative content of the sth element *by number* of atoms is

$$a_s = n_s/n_1 \,, \tag{5.1}$$

or, in the second case,

$$a_s(\%) = \frac{n_s}{n_1 + n_2 + \ldots + n_s + \ldots} \times 100 \,. \tag{5.2}$$

If the mass of one atom of the first element is m_1, the mass of one atom of the second m_2, etc., then the relative content of the sth element *by mass* is

$$\alpha_s = n_s m_s/n_1 m_1 \,, \tag{5.3}$$

or, as a percentage ratio,

$$\alpha_s(\%) = \frac{n_s m_s}{n_1 m_1 + n_2 m_2 + \ldots + n_s m_s + \ldots} \times 100 \,. \tag{5.4}$$

We must now make the following remark. In specifying for any star the quantities a_s (or α_s), we suppose that this chemical composition is *the same for all layers of its atmosphere*, i. e. the quantities a_s and α_s

do not change with depth in the photosphere. A number of facts indicate that this assumption is quite sound, i. e. that in stellar atmospheres there is *complete* mixing of the atoms of different elements. For instance, an analysis of the conditions in the solar corona has led I. S. Shklovskiĭ and other authors to the conclusion that the chemical composition of the solar corona is practically the same as the average composition of the chromosphere and of those layers of the solar atmosphere where absorption lines are formed (i. e. the photospheric layers). Since the extent of the layers of the solar corona where its chemical composition has been studied (the inner corona) is extremely great (about 200,000 km), it is in the highest degree improbable that any significant change in chemical composition could be observed inside such a thin layer as the Sun's photosphere (whose thickness is 100 to 300 km).

After this short introduction, we give a table which shows, for a series of celestial objects, the logarithms of the quantities n_s. The value of $\log_{10} n_1$ is arbitrarily taken in this table as 10 (9·8 for τ Scorpii).

Table 1

Element	Sun	τ Scorpii	Planetary nebulae	Interstellar gas
H	10·0	9·8	10·0	10·0
He	9·30(?)	9·1†	9·3	
C	6·89	6·0	6·1	
N	7·08	6·4	7·1	
O	6·65	6·8	7·3	
Na	4·33			4·4
Mg	5·57	5·6		
Ca	4·46			5·0
K	3·01			3·6
Al	4·17	4·3		
Si	5·12	5·6		
Fe	5·62*			
Cu	2·80			
Zn	2·52			
Sr	0·88			
Ba	0·38			
Hg	0·98			

* The results of different investigators regarding the content of iron are still quite considerably divergent.

† The most recent determinations give much smaller values of a_s for helium, of the order of 0·05; see Chapter 15.

A consideration of this table enables us to draw the following conclusions:

(1) The preponderant element in stellar atmospheres (and also in planetary nebulae and in the interstellar gas) is *hydrogen*. There are approximately ten thousand times as many hydrogen atoms in stellar

atmospheres as there are atoms of all the metals together. After hydrogen come helium, the atoms of the light elements C, N, O, and finally the metals.

(2) The chemical composition of the planetary nebulae and of the interstellar medium is practically the same as that of the solar and stellar atmospheres. This points to the identity of the chemical composition of the majority of celestial objects.

2. **The excitation and ionisation of atoms.** We now turn to the question of the *excitation and ionisation* of atoms. For the sake of clarity we shall use the terminology of Bohr's theory, and go over to quantum-mechanical concepts only where necessary.

Since it is necessary to introduce some terminology and notation, we recall briefly the basic postulates of this theory. In the absorption of a quantum $h\,\nu_{ik}$, an electron passes from an inner orbit i (a lower energy state of the atom) to an outer orbit k (a higher energy state). This process is called **excitation** of the atom. As a result of excitation, the internal energy of the atom increases by $h\,\nu_{ik}$, so that

$$\varepsilon_k - \varepsilon_i = h\,\nu_{ik}\,, \tag{5.5}$$

where ε_k and ε_i are the internal energies of the atom corresponding to the electron's being in the orbits k and i respectively.

As is customary, we shall use a convenient scheme of energy levels corresponding to the different possible states of the atom. Since we are interested only in the energy difference, we can ascribe an energy of zero to the lowest level. The energy necessary to transfer an electron from the lowest level to some given level is called the **excitation potential** of that level. Consequently, with the above-mentioned choice of the zero-point, the energies ε_i, ε_k, etc., which we have introduced are to be identified with the excitation potentials of the corresponding levels.

Hitherto we have spoken of the transition of an electron from an inner orbit to an outer one (the excitation of an atom). The reverse process, i. e. the jumping of an electron from an outer orbit k to an inner one i is accompanied by the *emission** of a quantum $h\,\nu_{ik}$ in accordance with the condition (5.5).

We now recall the basic postulates of the theory of ionisation of atoms. The separation of an electron from an atom (or ion) is called **ionisation** of the atom (or ion). As a result of single ionisation of a neutral atom, a positively charged ion and a free electron are formed. In double ionisation, the atom loses two electrons, and so on. The converse process to ionisation is called **recombination.** As a result of recombination, a positively charged ion reunites with a free electron. This leads to the

* Except in a collision of the second kind; see below.

formation of an ion whose degree of ionisation is one less than before the recombination. If the recombination takes place between a singly ionised atom and an electron, we obtain a neutral atom. The recombination of a neutral atom with an electron is also possible, and then a **negative ion** is formed.

A neutral atom of any element is distinguished by adding to the symbol of the element (on the right) the Roman figure I, for example: Ca I. For a singly ionised atom, two kinds of notation can be used: (1) a plus sign is placed above and to the right of the symbol of a chemical element; (2) the Roman figure II is placed to the right of the symbol; for example: Ca^+ or Ca II. In the case of a doubly ionised atom of any element, two plus signs are placed above the symbol of the element, or the Roman figure III to the right of it; for example: Ca^{++} or Ca III. This symbolism is continued for higher states of ionisation. The second method of writing the ion (with Roman figures) is in some cases to be preferred to the first, which is very inconvenient for multiple ionisation.

The detachment of an optical electron from an atom (or ion) can take place, in general, from any orbit. The energy required to detach an electron from the orbit with $\varepsilon = 0$, without giving any kinetic energy to the detached electron, is called the **ionisation potential** of the atom. This potential will be denoted by χ_r. In the case of a neutral atom $r = 0$, for a singly ionised atom $r = 1$, and so on.

In order to know whether the neutral or the ionised atom is being discussed, it is convenient to add to the excitation potential ε_k a symbol showing the degree of ionisation of the atom in question. In other words, we write for instance

$$\varepsilon_{r,\,k}\,.$$

The energy necessary to ionise the atom from any excited level is called the **binding energy** of this level. We shall denote it by $\chi_{r,\,k}$. It is clear that we always have

$$\chi_r = \chi_{r,\,k} + \varepsilon_{r,\,k}\,. \qquad (5.6)$$

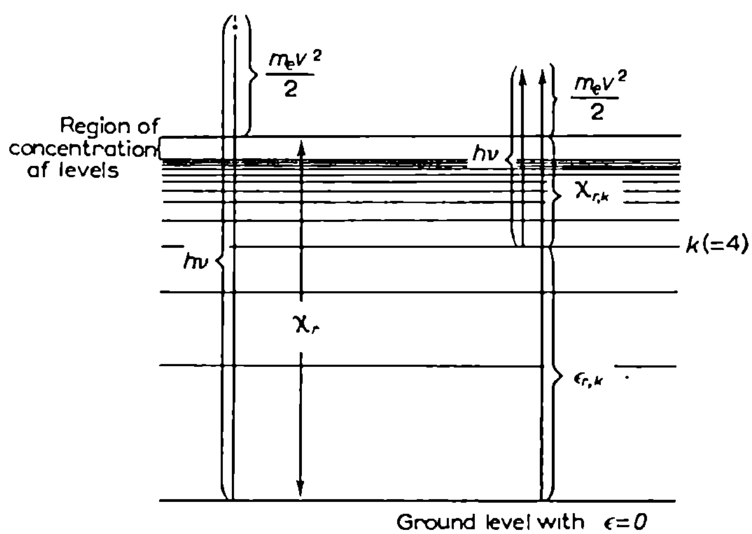

FIG. 10

The scheme of energy levels according to our notation is represented in Fig. 10. We deliberately do not assign any affix to the lowest level, since the quantum numbers of this level may in general vary greatly.

On the left of Fig. 10 is shown the case where the magnitude of the quantum $h\nu$ absorbed by the atom is greater than χ_r, i. e. the atom is *ionised by radiation*. In this case, besides the detachment of the electron, a velocity v is given to the latter, and it moves with kinetic energy $\frac{1}{2} m_e v^2$, where the fundamental equation of the photo-effect

$$h\nu = \chi_r + \tfrac{1}{2} m_e v^2 \tag{5.7}$$

holds, m_e being the mass of the electron.

If the ionisation takes place from some excited level (i. e. a level with an energy greater than that of the lowest level, where $\varepsilon \equiv 0$), then instead of (5.7) we have

$$h\nu = \chi_{r,k} + \tfrac{1}{2} m_e v^2 . \tag{5.8}$$

[This case is shown on the right of Fig. 10 — *Translator.*] In recombination the electron may fall into any level, and as a result a quantum $h\nu$ is emitted, the equation (5.7) or (5.8) being satisfied.

Let us now consider the quantitative formulae describing the excitation and ionisation of atoms in a gas. In doing so we shall for the present use formulae which are valid for *thermodynamic equilibrium*. The question of how these formulae must be modified for application to the conditions existing in stellar photospheres will be considered later.

We first examine the formula which describes the excitation of atoms. Let the number of r times ionised atoms in the excitation state k contained in 1 cm^3 be $n_{r,k}$. Also, let the number of these atoms in the excitation state i contained in the same volume be $n_{r,i}$. In thermodynamic equilibrium, Boltzmann's formula holds:

$$\frac{n_{r,k}}{n_{r,i}} = \frac{g_{r,k}}{g_{r,i}}\, e^{-(\varepsilon_{r,k} - \varepsilon_{r,i})/kT} , \tag{5.9}$$

where $g_{r,k}$ and $g_{r,i}$ are the statistical weights* of the levels k and i, k is Boltzmann's constant and T is the absolute temperature. If the lower state i is the lowest state, then $\varepsilon_i \equiv 0$ and only ε_k remains in the exponent.

In cases where only r times ionised atoms are considered, r being fixed, the suffix r may be omitted. In this case

$$\frac{n_k}{n_i} = \frac{g_k}{g_i}\, e^{-(\varepsilon_k - \varepsilon_i)/kT} . \tag{5.10}$$

* The *statistical weight* (or degree of degeneracy) is the name given to the number of different states of the quantum system (the atom) corresponding to any energy level ε_i. It shows into how many terms the state considered would split if the degeneracy were completely removed (for instance, by means of an external magnetic field) or, expressed in the more rigorous language of quantum mechanics, how many linearly independent eigenstates belong to the energy value considered. For hydrogen, the statistical weight is $2n^2$, where n is the principal quantum number.

The statistical weights and the energies $\varepsilon_{r,k}$ are taken from the theory of spectra.

The following important remark must be made in connection with formulae (5.9) and (5.10). Since the difference between the energy of any excited level and the energy of the *ground* level is usually large, it follows from the two formulae mentioned that in such cases the majority of atoms are in the ground (non-excited) state. Thus, for example, the ratio of the number of hydrogen atoms in the second state to the number in the ground state at a temperature $T = 5700°$ is $4{\cdot}2 \times 10^{-9}$.

We now turn to the formula for ionisation. Let the number of r times ionised atoms, in all possible excited states, contained in 1 cm^3 be n_r (i. e. $n_r = \sum_k n_{r,k}$). Also, let the total number of $r+1$ times ionised atoms in the same volume be n_{r+1}. Then in *thermodynamic equilibrium* Saha's ionisation formula holds [168, p. 83]:

$$\frac{n_{r+1}}{n_r}\, p_e = \frac{u_{r+1}}{u_r}\, \frac{2(2\pi m_e)^{\frac{3}{2}} (kT)^{\frac{5}{2}}}{h^3}\, e^{-\chi_r/kT} \,. \tag{5.11}$$

To calculate the ionisation, it is more convenient to use Saha's formula in the logarithmic form

$$\log_{10} \frac{n_{r+1}}{n_r} = \log_{10} \frac{u_{r+1}}{u_r} + \tfrac{5}{2} \log_{10} T - \frac{5040}{T}\, \chi_r - \\ - \log_{10} p_e + \log_{10} \frac{2(2\pi m_e)^{\frac{3}{2}} k^{\frac{5}{2}}}{h^3} \,. \tag{5.11 a}$$

Here χ_r is expressed in electron-volts, and not in ergs as in formula (5.11).

In formula (5.11), p_e is the partial electron pressure. If there are n_e *free* electrons in 1 cm^3, then

$$p_e = n_e\, kT \,. \tag{5.12}$$

The quantity u_r is what is called the **partition function** of r times ionised atoms. If we allot the suffix 1 to the ground state (for which $\varepsilon \equiv 0$) and begin the numbering of the successive levels from there, then u_r has the form

$$u_r(T) = g_{r,1} + g_{r,2}\, e^{-\varepsilon_{r,2}/kT} + g_{r,3}\, e^{-\varepsilon_{r,3}/kT} + \\ + \ldots = \sum_{k=1}^{\infty} g_{r,k}\, e^{-\varepsilon_{r,k}/kT} \,, \tag{5.13}$$

where $g_{r,k}$ are the statistical weights of the successive levels of the r times ionised atom. Very often the inequalities

$$e^{-\varepsilon_{r,2}/kT}\,,\ e^{-\varepsilon_{r,3}/kT}\,,\ \ldots \ll 1 \tag{5.14}$$

hold. In this case the function $u_r(T)$ reduces to the statistical weight of the *ground* level. If, however, the first excited level is less than 1 eV from the ground level (1 electron-volt $= 1{\cdot}602 \times 10^{-12}$ erg), the second term of the sum (5.13) must also be taken into account, and sometimes several terms. If, on the other hand, the excited level is 2 eV from the ground level, then it is necessary to take account of the second term of (5.13) only for fairly high temperatures ($T > 10{,}000°$)*.

The quantity $u_{r+1}(T)$ denotes the same function of the temperature as $u_r(T)$, but for the $r+1$ times ionised atom. The factor $2u_{r+1}/u_r$ is generally of the order of unity. A table of the ionisation potentials for various elements in various stages of ionisation, together with nomograms for the employment of Saha's formula, is contained in A. UNSÖLD's book [168, § 23].

In connection with (5.11) we can introduce what is called the **degree of ionisation** x_r, which we define as follows:

$$x_r = \frac{n_r}{n_0 + n_1 + \ldots + n_r + \ldots} \quad (r=0,1,2,\ldots)\,. \tag{5.15}$$

Thus the degree of ionisation gives the proportion of r times ionised atoms among the total number of atoms of the given kind.

It follows immediately from Saha's formula (5.11) that for a *fixed* p_e *the degree of ionisation of any element increases monotonically with the temperature* T, and this increase is proportionately the more rapid, the greater the ionisation potential χ_r and the lower the temperature T. For small χ_r/kT, the ionisation increases comparatively slowly with T. It also follows from this formula that *when* p_e *increases* (for a fixed T) *the ionisation decreases.*

In the majority of cases encountered in practice, the levels of successive ionisation states are so far apart that only atoms in two neighbouring ionisation states, r and $r+1$, can simultaneously exist in fair quantity in a given volume in which p_e is fixed. For instance, if neutral and singly ionised atoms exist in an appreciable number, then the number of doubly ionised atoms will be negligibly small, and so on. Taking into account only two successive ionisation states, r and $r+1$, we can say that, if some proportion x of the atoms are $r+1$ times ionised, then the proportion $1-x$ will be r times ionised. Consequently, formula (5.11) can be rewritten as

$$\frac{x}{1-x}\, p_e = K_r\,, \tag{5.16}$$

* For a calculation of the sums u_0 and u_1 at temperatures from 4900° to 7300° and for various elements, see W. J. CLAAS [33]. The convergence of sums of the form (5.13) is discussed in A. UNSÖLD's book [168, p. 84].

where

$$K_r = \frac{u_{r+1}}{u_r} \frac{2(2\pi m_e)^{\frac{3}{2}} (kT)^{\frac{5}{2}}}{h^3} e^{-\chi_r/kT} , \tag{5.17}$$

and also

$$x = n_{r+1}/n , \quad 1 - x \approx n_r/n . \tag{5.18}$$

In conclusion, we shall indicate an important modification of formula (5.11). Very often the number of atoms in the ground state is determined from observation (see Part II). For this case, formula (5.11) takes the form

$$\frac{n_{r+1,1}}{n_{r,1}} p_e = \frac{g_{r+1,1}}{g_{r,1}} \frac{2(2\pi m_e)^{\frac{3}{2}} (kT)^{\frac{5}{2}}}{h^3} e^{-\chi_r/kT} \tag{5.19}$$

The second suffix shows that the $r + 1$ times and r times ionised atoms are in the ground state.

3. The true absorption of radiation. Let us now pass to the question of the absorption coefficient. It is necessary to make some introductory remarks.

The temperature distribution in any region of a stellar photosphere is determined, as we have seen in the preceding chapters, by the interaction between *radiation* and *matter*. In particular, the processes of thermal conduction and convection play only a subordinate part in establishing the temperature distribution. Practically all the *thermal* energy which is gained each second by any element of matter in the photosphere comes from the absorption of radiation. In this respect we shall at present be interested only in those absorption processes which lead to the conversion of radiant energy into thermal energy, and conversely, of thermal into radiant energy. The mechanism of the passage of energy outwards from the interior parts of the star amounts largely to the operation of these processes.

It is implied with regard to processes of this kind that *there is no direct connection* between the absorbed and emitted quanta. Each absorbed quantum of frequency ν is entirely lost, and the thermal energy thereby gained by the material of the photosphere is emitted in other frequencies after some time. All this is reflected in the condition of radiative equilibrium (3.14), which asserts that the total amount of energy absorbed in one second by unit volume (or mass) equals the total amount of energy emitted by that volume in the same time.

Absorption in which the energy absorbed is converted into thermal energy (with subsequent re-emission in other frequencies) we shall call **true absorption** (as distinct from scattering, of which we shall speak later). In order to make clearer what has been explained above, we give two examples of true absorption.

(1) Suppose that photo-ionisation has taken place from some orbit k as the result of the absorption of a quantum $h\nu$ by the atom. The detached electron thereby acquires a velocity v in accordance with formula (5.8). Having undergone a great number of collisions with other particles (chiefly with electrons; see Chapter 8), it will finally be captured by some ion, i. e. recombination will take place. The frequency of the quantum then emitted will, in general, be different from the frequency of the absorbed quantum, since the velocity of the electron at the moment of recombination will be different from its original velocity v at the moment of photo-ionisation. The electron may be captured by a completely different ion, or into a different orbit of the same ion. At the same time, there is equality between the absorbed and emitted energy for the whole spectrum, as has already been stated.

(2) A second example is *true selective* absorption (i. e. absorption in the frequency of spectral lines). Suppose that, as a result of the absorption of a quantum $h\,\nu_{ik}$, an electron has passed from the orbit i to an outer orbit k, and that at the moment when it is in the latter orbit the atom undergoes an inelastic collision (a collision of the second kind) with some other particle. As a result of the collision the excitation energy of the atom (i. e. $\varepsilon_k - \varepsilon_i$) may be transformed into kinetic energy of the particle. In this way the energy $h\,\nu_{ik}$ of the quantum is entirely converted into the energy of thermal motion of the particle, and the electron falls back to the lower orbit i, but the atom does not emit a quantum $h\,\nu_{ik}$. Another variant of the case considered is also possible. The electron may be removed by photo-ionisation from the excited level k. Here too the energy of the absorbed quantum is converted into energy of thermal motion. However, it must be admitted that, for the temperature problem in which we are interested, both variants of the present example, which have to do only with *selective absorption and emission in a definite frequency*, are of subordinate interest: (1) calculations which are performed in Part II indicate that processes of true selective absorption generally play a relatively minor part* in the total energy balance; (2) the total interval of the spectrum in which these processes are noticeable is generally very small. We shall therefore consider henceforward only processes of true absorption for the continuous spectrum. Such processes will be called **true continuous absorption** or **general true absorption.** For the sake of brevity the word *true* can be omitted, as is usually done, and we speak simply of **continuous** or **general absorption.** We shall begin our discussion with processes of photoelectric absorption, which belongs to the category of true absorption. This absorption process plays the most important part in stellar atmospheres.

* For instance, collisions of the second kind are rare, on account of the small density of matter in stellar atmospheres.

Let us consider the process of the photo-ionisation of an atom from some level k. It follows from (5.8) that the ionisation of the atom from this level can be brought about by any quantum, provided that the energy of the latter is greater than or equal to the binding energy $\chi_{r,k}$. Thus, if continuous radiation falls on an assembly of atoms, then this radiation, by ionising atoms from the level k, will be absorbed in the frequency interval from $\nu_k = \chi_{r,k}/h$ to $\nu = \infty$. In other words, a sufficiently thick absorbing layer (i. e. one of large optical thickness) will produce in the continuous radiation passing through it a continuous absorption band beginning at the frequency ν_k and going in the direction of greater frequencies. And since each atom has an infinite number of possible orbits, the photoelectric absorption must in general result in an infinite sequence of absorption bands. In some cases (for example, the negative hydrogen ion) there is only a finite number of absorption bands (see below).

In order to describe such an absorption process quantitatively, we first consider a simple system, namely a hydrogen or hydrogen-like ion. (A hydrogen-like ion consists of a nucleus and one electron. To this class belong He II, Li III, Be IV, etc.)

Both for the following discussion and for a number of astrophysical problems it is more convenient to calculate the absorption not for unit mass of absorbing material, as for instance in (3.13), but for one absorbing atom. Let the mass of this absorbing atom, having an absorption coefficient $\varkappa_\nu$, be m. Then

$$\varrho = mn\,, \tag{5.20}$$

where n is the number of absorbing atoms of the given kind in 1 cm^3, and ϱ, as in (3.13), is the density of these atoms. Introducing (5.20) on the right of (2.13), we obtain

$$\mathrm{d}I_\nu = -\, I_\nu\, \varkappa_\nu\, mn\, \mathrm{d}h\,. \tag{5.21}$$

We also write

$$k_\nu = \varkappa_\nu\, m\,. \tag{5.22}$$

Then

$$\mathrm{d}I_\nu = -\, I_\nu\, k_\nu\, n\, \mathrm{d}h\,. \tag{5.23}$$

We shall call k_ν the **absorption coefficient referred to one atom** (often called simply the **atomic absorption coefficient**). The dimensions of k_ν are cm^2. The expression (5.23) can be written down at once without using (2.13). However, the intermediate steps are necessary to obtain the formula (5.22) which connects the absorption coefficient referred to unit mass with that referred to one atom.

4. **Photoelectric absorption for hydrogen-like atoms.** We consider a hydrogen-like atom with an electron moving in the field of the charge Ze, in an orbit with principal quantum number n. In this case, quantum mechanics gives for k'_ν the following expression*:

$$k'_{\nu(n,\,\Sigma\, l)} = \frac{32\,\pi^2\, e^6 R\, Z^4}{3\sqrt{3}\, ch^3\, \nu^3\, n^5}\, g'\,, \tag{5.24}$$

where the suffix $\Sigma\, l$ indicates that the coefficient of photoelectric absorption is calculated by taking account of *all* the sub-levels which belong to the given principal quantum number n, R is Rydberg's frequency

$$R = 2\pi^2\, e^4\, m_e/h^3\,, \tag{5.25}$$

which equals $3{\cdot}287871 \times 10^{15}$ sec^{-1}, and g' is some correction factor little different from unity. According to D. H. Menzel and C. L. Pekeris [86] it is

$$g' \approx 1 - 0{\cdot}1728 \left(\frac{\nu}{R\, Z^2}\right)^{\frac{1}{3}} \left[\frac{2}{n^2}\left(\frac{R\, Z^2}{\nu}\right) - 1\right]. \tag{5.26}$$

It follows from (5.24) that within each band, beginning at the frequency $\nu_n = \chi_n/h$ and going to $\nu = \infty$, the absorption coefficient decreases as $1/\nu^3$. Calculations give the numerical values shown in Table 2 for k'_ν for hydrogen ($Z = 1$) at the frequencies corresponding to the edges of the bands.

Table 2

Initial state	$k_\nu{}'_{(n,\,l)}$ (cm^2)
1 *s*	$6{\cdot}3 \times 10^{-18}$
2 *s*	$1{\cdot}5 \times 10^{-17}$
2 *p*	$1{\cdot}4 \times 10^{-17}$
3 *s*	$2{\cdot}6 \times 10^{-17}$
3 *p*	$2{\cdot}6 \times 10^{-17}$
3 *d*	$1{\cdot}8 \times 10^{-17}$

Here the values of k'_ν are given separately for different azimuthal numbers l.

The absorption coefficient given by formula (5.24) has been derived for an atom in the quantum state n, and naturally does not depend on any variable parameters such as, for example, pressure, temperature, etc. However, this assertion is justified only while we do not consider all absorption bands. In fact, photoelectric ionisation (and also absorption

* The prime added to k_ν signifies that no account has yet been taken of the stimulated emission; see section 5, equation (5.61) — *Translator.*

of radiation) can take place from any orbit of the atom. Let us consider any definite frequency ν. From this frequency in the direction of smaller frequencies (i. e. to the long-wave end of the spectrum) lie all band edges at frequencies $\nu_n = \chi_n/h \leq \nu$. And since the number of orbits in the atom is infinite, the number of these bands is also infinite. Every such band will contribute to the absorption at frequency ν, since the absorption bands go to $\nu = \infty$. Fig. 11 illustrates what we have said. In this

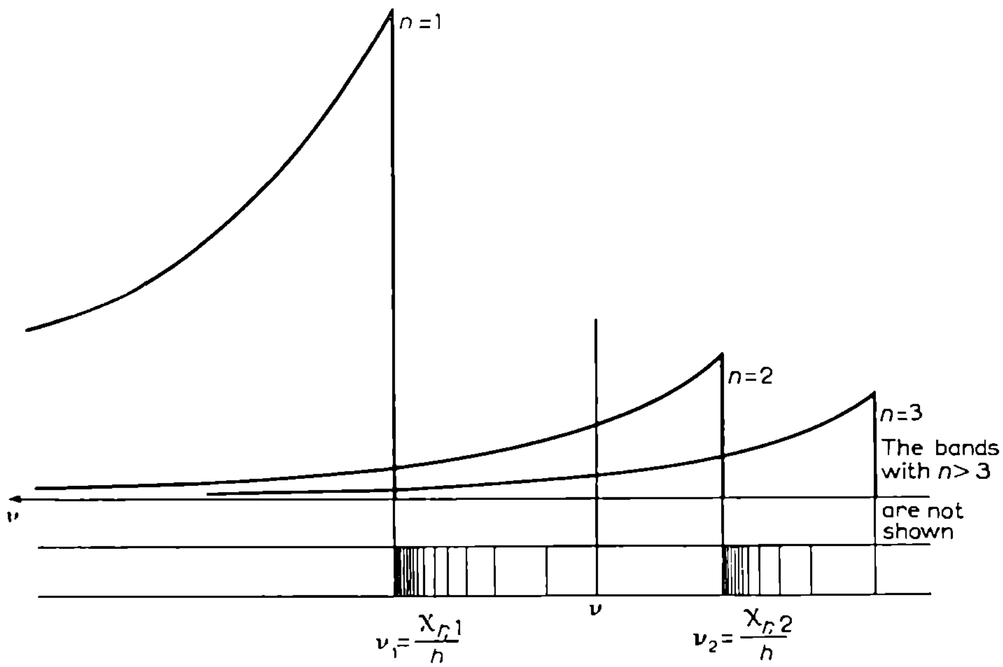

FIG. 11

figure the frequency ν considered lies between ν_1 and ν_2. The bands with $n = 2, 3, 4, \ldots, \infty$ correspond to the photo-ionisation of the atom from all higher levels. (In accordance with Fig. 10 the binding energy χ_n, and therefore the frequency $\nu_n = \chi_n/h$, diminishes as n increases.)

Further, as follows from very general properties of spectral series, the frequency of the limit of the series arising from the orbit of principal quantum number n is determined by the binding energy of the level in question, i. e. by χ_n. The bands of photo electric absorption start from the limits of the corresponding series, and this also is shown in Fig. 11 (see the portion of the spectrum at the bottom, where the series are indicated).

Let us now return to the topic which we were considering. We have said that all those bands contribute to the absorption at frequency ν, the frequencies of whose edges are less than the given frequency. Consequently, in order to obtain the resultant absorption coefficient for the frequency in question, we must take account of all these bands, and to do this it is necessary to make some hypothesis regarding the distribu-

tion of atoms among the quantum states. We shall assume that the distribution of atoms among the states is determined by Boltzmann's formula (5.9), which we shall rewrite

$$\frac{n_{r,n}}{n_{r,1}} = \frac{g_{r,n}}{g_{r,1}} e^{-\varepsilon_{r,n}/kT}, \tag{5.27}$$

where we take $\varepsilon_{r,1} \equiv 0$. In virtue of (5.27), (5.23) and (5.24) the absorbing effect of the atoms at frequency ν due only to the nth absorption band is

$$\mathrm{d}I_\nu^{(n)} = - I_\nu k'_{\nu(n,\Sigma l)} n_{r,n} \mathrm{d}h = - I_\nu \frac{32\pi^2 e^6 R Z^4 g' n_{r,1} g_{r,n}}{3\sqrt{3}\, c h^3 \nu^3 n^5 g_{r,1}} e^{-\varepsilon_{r,n}/kT} \mathrm{d}h\,. \tag{5.28}$$

And since at frequency ν the absorption is produced by all the bands, the frequencies of whose edges are less than ν, the total absorption is given by the formula

$$\mathrm{d}I_\nu = \sum_{n_0}^{\infty} \mathrm{d}I_\nu^{(n)} = - I_\nu \frac{32\pi^2 e^6 R Z^4 n_{r,1}}{3\sqrt{3}\, c h^3 \nu^3 g_{r,1}} \sum_{n_0}^{\infty} \frac{g_{r,n}}{n^5} g' e^{-\varepsilon_{r,n}/kT} \mathrm{d}h\,, \tag{5.29}$$

where the sum begins at the first band which lies in the direction of smaller frequencies from ν (see Fig. 11), this band belonging to the principal quantum number n_0.

In connection with formula (5.29) it must be pointed out that the intensity of the successive bands must *decrease* as the principal quantum number increases, i. e. in the direction of smaller frequencies. This is chiefly due to the presence in the formula of the Boltzmann factor $e^{-\varepsilon_{r,n}/kT}$. The larger $\varepsilon_{r,n}$, the fewer the atoms corresponding to this value of $\varepsilon_{r,n}$ and the weaker the absorption of the gas.

In order to pass from $n_{r,1}$ to n, the total number of the atoms in question in 1 cm^3, $n_{r,n}$ from (5.27) must be summed over all states n from $n = 1$ to $n = \infty$, taking account of (5.13):

$$n_r = \sum_{n=1}^{\infty} n_{r,n} = \frac{n_{r,1}}{g_{r,1}} \sum_{n=1}^{\infty} g_{r,n} e^{-\varepsilon_{r,n}/kT} = \frac{n_{r,1} u_r}{g_{r,1}}\,. \tag{5.30}$$

By virtue of (5.30) and the ionisation equation (5.11), we obtain

$$\frac{n_{r,1}}{g_{r,1}} = \frac{n_{r+1}}{u_{r+1}} p_e \frac{h^3 e^{\chi_r/kT}}{2(2\pi m_e)^{\frac{3}{2}} (kT)^{\frac{5}{2}}}\,. \tag{5.31}$$

Introducing now (5.31) into (5.29), we obtain

$$\mathrm{d}I_\nu = \sum_{n_0}^{\infty} \mathrm{d}I_\nu^{(n)} =$$

$$= - I_\nu \frac{n_{r+1} p_e}{u_{r+1}} \frac{2^4 \pi^2 e^6 R Z^4 \mathrm{d}h}{3\sqrt{3}\, c (2\pi m_e)^{\frac{3}{2}} (kT)^{\frac{5}{2}} \nu^3} \sum_{n_0}^{\infty} \frac{g_{r,n} g'}{n^5} e^{-(\varepsilon_{r,n} - \chi_r)/kT}\,. \tag{5.32}$$

For the hydrogen-like atoms considered, the well-known series formula

$$\varepsilon_{r,\,n} = h\,R\,Z^2\,(1 - 1/n^2) = \chi_r - h\,R\,Z^2/n^2 \tag{5.33}$$

holds. For these atoms, $u_{r+1} = 1$, since in the case in question the $r+1$ times ionised atom is the *nucleus* with charge $+Ze$, and consequently the partition function reduces to one term, the statistical weight of the nucleus. Taking these facts into account and noticing that from the theory of spectra $g_{r,\,n} = 2n^2$, we obtain

$$\mathrm{d}I_\nu = - I_\nu\, n_{r+1}\, p_e \frac{C_0 Z^2 T^{-\frac{3}{2}}}{\nu^3} \left[\frac{2\,h\,R\,Z^2}{kT} \sum_{n_0}^{\infty} \frac{g'}{n^3}\, e^{h R Z^2/n^2 kT}\right] \mathrm{d}h\,, \tag{5.34}$$

where

$$C_0 = \frac{2^4\,\pi^2\,e^6}{3\sqrt{3}\,c\,h}\,\frac{1}{(2\,\pi\,m_e\,k)^{\frac{3}{2}}} = 2{\cdot}67 \times 10^{24}\,. \tag{5.35}$$

5. **The allowance for stimulated emission.** The expression (5.34) is not the final one. We must allow also for **stimulated emission**, or, as it is sometimes called, **negative absorption.** The processes of interaction between radiation and matter can be described not only by means of the emission and absorption coefficients, but also by means of the coefficients of transition probability. Let the number of atoms of any material in the quantum state k contained in 1 cm^3 be n_k, and the number of such atoms in the lower state i be n_i. Then the number of transitions in 1 cm^3 of material in one second which are accompanied by the emission of a quantum $h\nu_{ik}$ is

$$n_{k\to i} = n_k (A_{ki} + \varrho_{\nu_{ik}}\, B_{ki})\,, \tag{5.36}$$

where

$$\varrho_{\nu_{ik}} = \frac{1}{c} \int I_{\nu_{ik}}\,\mathrm{d}\omega \tag{5.37}$$

is the radiation density* in the given frequency, and A_{ki} and B_{ki} are the transition probability coefficients; A_{ki} corresponds to *spontaneous* transitions, and $\varrho_{\nu_{ik}}\, B_{ki}$ to *stimulated* transitions which take place under the action of the radiation field of density $\varrho_{\nu_{ik}}$.

* We can obtain formula (5.37) from the calculations which we performed at the beginning of Chapter 3. According to (3.3) and Fig. 5, the quantity of energy passing in one second through the element $\mathrm{d}\sigma$ inside the solid angle $\mathrm{d}\omega$ and in unit frequency interval is $I_\nu\,\mathrm{d}\sigma\cos\theta\,\mathrm{d}\omega$. The time necessary for a beam to pass through the height $\mathrm{d}s$ of the cylinder $\mathrm{d}v$ is $\mathrm{d}s/c$. Consequently, the amount of energy from the beam $I_\nu\,\mathrm{d}\omega$ considered which is contained in the cylinder $\mathrm{d}v$ at any given moment is $I_\nu\,\mathrm{d}\sigma\cos\theta\,\mathrm{d}\omega.\,\mathrm{d}s/c = I_\nu\,\mathrm{d}\omega\,\mathrm{d}v/c$. Integrating this equation over the whole sphere and the whole volume Δv and dividing the result by Δv, we obtain the required energy density, i. e. formula (5.37).

The number of converse transitions, i. e. transitions $i \to k$, which take place in 1 cm^3 in one second under the action of the radiation field and are accompanied by the absorption of quanta $h\nu_{ik}$ is

$$n_{i\to k} = n_i \, \varrho_{\nu_{ik}} \, B_{ik} \, , \tag{5.38}$$

where B_{ik} is again the corresponding coefficient of transition probability.

In using (5.36) and (5.38) it is assumed that the radiation density does not vary very greatly in the immediate neighbourhood of ν_{ik}, i. e. within the spectral line considered, which is produced by the transitions $i \to k$ and $k \to i$.

The following relations exist between the transition coefficients A_{ki}, B_{ki} and B_{ik}:

$$g_k \, B_{ki} = g_i \, B_{ik} \, , \tag{5.39}$$

$$A_{ki} = \frac{8\pi h \nu_{ik}^3}{c^3} B_{ki} = \frac{8\pi h \nu_{ik}^3}{c^3} \frac{g_i}{g_k} B_{ik} \, . \tag{5.40}$$

Using (5.40), we can rewrite (5.36) as

$$n_{k\to i} = n_k \, A_{ki} \left(1 + \frac{c^3}{8\pi h\nu_{ik}^3} \, \varrho_{\nu_{ik}}\right) . \tag{5.41}$$

The importance of the stimulated emission, as compared with the spontaneous emission, is determined by the second term in parentheses in (5.41).

The expressions (5.36), (5.38) and (5.41) are written for the integrated density (5.37) (i. e. taking account of all directions). For a *narrow* beam of rays of dispersion dω, the density of radiant energy is

$$(\varrho_{\nu_{ik}})_{\mathrm{d}\omega} = I_{\nu_{ik}}(0) \, \mathrm{d}\omega / c \, , \tag{5.42}$$

and the expressions (5.36) and (5.38) become

$$(n_{k\to i})_{\mathrm{d}\omega} = n_k \left(\frac{A_{ki}}{4\pi} + \frac{I_{\nu_{ik}}(0)}{c} B_{ki}\right) \mathrm{d}\omega \, , \tag{5.43}$$

$$(n_{i\to k})_{\mathrm{d}\omega} = n_i \, \frac{I_{\nu_{ik}}(0)}{c} \, B_{ik} \, \mathrm{d}\omega \, , \tag{5.44}$$

while (5.43), on using (5.40), can be written

$$(n_{k\to i})_{\mathrm{d}\omega} = n_k \, A_{ki} \left(\frac{1}{4\pi} + \frac{c^3}{8\pi h\nu_{ik}^3} \frac{I_{\nu_{ik}}(0)}{c}\right) \mathrm{d}\omega \, . \tag{5.45}$$

The expression (5.34) for the attenuation of a beam of radiation takes account only of the ordinary absorption, to which the coefficient B_{ik} corresponds. However, the appropriate analysis shows that, in the pres-

ence of stimulated emission processes, the use of the basic equation of transfer (3.12) requires a modification of the expression (5.34). For this we have to introduce the transition coefficients for recombination processes. Let k'_ν be the atomic absorption coefficient corresponding to the photo-ionisation, from the kth level, of an r times ionised atom. Then, according to (3.6), the number of photo-ionisation processes from level k which take place in 1 cm^3 during one second and are accompanied by the absorption of quanta with frequencies between ν and $\nu + \mathrm{d}\nu$ is

$$n_{k\to e}\,\mathrm{d}\nu = n_{r,\,k}\,k'_\nu\,\frac{\mathrm{d}\nu}{h\nu}\int I_\nu\,\mathrm{d}\omega\,, \tag{5.46}$$

since in every photo-ionisation a quantum $h\nu$ is absorbed. In every such process the final state of the electron is the free state with some definite velocity v, determined by the equation of the photoelectric effect. It is to be noticed that we have passed, in accordance with (5.20) and (5.22), from the absorption coefficient referred to unit mass to the atomic absorption coefficient.

The expression (5.46) may also be rewritten, in accordance with (5.37),

$$n_{k\to e}\,\mathrm{d}\nu = n_{r,\,k}\,k'_\nu\,(c\,\mathrm{d}\nu/h\nu)\,\varrho_\nu\,. \tag{5.47}$$

The converse processes to the photo-ionisation of the atom from the level k are the recombinations to this level. The process of recombination may be regarded as a collision of the ion with an electron. Hence we can introduce the "effective cross-section" for recombination, which we denote by β_ν. Then the number of spontaneous recombinations taking place during one second in 1 cm^3 and accompanied by the emission of quanta with frequencies between ν and $\nu + \mathrm{d}\nu$ is

$$n'_{e\to k}\,\mathrm{d}\nu = n_{r+1}\,\beta_\nu\,v\,\mathrm{d}n_e\,, \tag{5.48}$$

where $\mathrm{d}n_e$ is the number of electrons contained in 1 cm^3 and having velocities between v and $v + \mathrm{d}v$, where $\mathrm{d}v$ and $\mathrm{d}\nu$ are connected, in accordance with (5.8), by the relation

$$v\,\mathrm{d}v = (h/m_e)\,\mathrm{d}\nu\,. \tag{5.49}$$

The form of the expression (5.48) is obtained as follows. Let λ_e be the mean free path of an electron (between atoms of the kind in question). Then an electron having velocity v undergoes v/λ_e collisions in 1 second. Consequently the number of collisions $\mathrm{d}Z$ undergone by $\mathrm{d}n_e$ electrons in 1 second is $(v/\lambda_e)\,\mathrm{d}n_e$. On the other hand, in collisions of electrons with atoms and ions, both the latter may be regarded as at rest; in this case λ_e, by the kinetic theory of gases, is equal to $1/q\,n$, where q is the effective cross-section for collisions, and n is the number of atoms

(or ions) in 1 cm^3 with which electrons are colliding (i. e. recombining). Substituting the expression for λ_e into that for dZ, we obtain an expression of the form (5.48).

The expression (5.48) takes account only of the processes of *spontaneous recombination*. However, if the medium is in a field of radiation of density ϱ_ν, we must also take account of the processes of stimulated recombination, which gives from (5.41)

$$n_{e\to k}\,d\nu = n_{r+1}\,\beta_\nu\,v\left(1+\frac{c^3}{8\pi h\nu^3}\varrho_\nu\right)dn_e\,. \tag{5.50}$$

In order to find the connection between β_ν and k'_ν, we make use of the **principle of detailed balancing**, which asserts that *in a state of thermodynamic equilibrium any process takes place exactly as often as the converse process*. Consequently, in thermodynamic equilibrium the quantities (5.47) and (5.50) must be equal, and so

$$n_{r,k}\,k'_\nu\,\frac{c\,d\nu}{h\,\nu}\,\varrho_\nu = n_{r+1}\,\beta_\nu\,v\left(1+\frac{c^3}{8\pi h\nu^3}\varrho_\nu\right)dn_e\,. \tag{5.51}$$

Next, in the presence of thermodynamic equilibrium, there exists a Maxwellian distribution of particles among velocities:

$$dn_e = n_e\,4\pi\left(\frac{m_e}{2\pi kT}\right)^{\frac{3}{2}} e^{-m_e v^2/2kT}\,v^2\,dv\,. \tag{5.52}$$

For ϱ_ν, in the presence of thermodynamic equilibrium, when $I_\nu(\theta)\equiv B_\nu(T)$, we have, according to (5.37) and (3.11),

$$\varrho_\nu = \frac{8\pi h\nu^3}{c^3}\,\frac{1}{\exp(h\nu/kT)-1}\,. \tag{5.53}$$

Finally, by (5.9) and (5.30) we have for $n_{r,k}$

$$n_{r,k} = n_{r,1}\,\frac{g_{r,k}}{g_{r,1}}\,e^{-\varepsilon_{r,k}/kT} = \frac{n_r}{u_r}\,g_{r,k}\,e^{-\varepsilon_{r,k}/kT}\,, \tag{5.54}$$

while to obtain the connection between n_r, n_{r+1} and n_e we can use Saha's formula (5.11), taking account of (5.12).

Introducing all these results into (5.51), and using (5.49), (5.6) and (5.8), we find the required expression:

$$\beta_\nu = \frac{g_{r,k}}{u_{r+1}}\,\frac{\nu^2 h^2}{c^2 m_e^2 v^2}\,k'_\nu\,. \tag{5.55}$$

We have derived the expression (5.55) for the case of thermodynamic equilibrium. However, it is completely general, since it represents the relation between quantities appertaining to the atoms and ions themselves. We recall that the analogous expressions (5.39) and (5.40) for

discrete transitions were also obtained for thermodynamic equilibrium. They are untrue only in the case where the radiation field varies very rapidly with frequency near ν_{ik}.

Let us now consider the original equation of transfer (2.28), in which we shall understand by $\varkappa_\nu$ the absorption coefficient which takes account only of ordinary absorption. We mark this coefficient with a prime, and we then have

$$\cos\theta\,\frac{\mathrm{d}I_\nu(\theta)}{\mathrm{d}h}\,\mathrm{d}\nu\,\mathrm{d}\omega = I_\nu(\theta)\,\varkappa'_\nu\,\varrho\,\mathrm{d}\nu\,\mathrm{d}\omega - j_\nu(\theta)\,\varrho\,\mathrm{d}\nu\,\mathrm{d}\omega\,. \qquad (5.56)$$

We consider the equation (5.56) for processes connected with the ionisation of the atom from the level k and with recombination to this level. The quantity $j_\nu(\theta)\,\mathrm{d}\nu\,\mathrm{d}\omega$ is just the energy emitted by one gram of matter in the photosphere in one second, in the frequency interval from ν to $\nu + \mathrm{d}\nu$ and within the solid angle $\mathrm{d}\omega$. The dependence of $j_\nu(\theta)$ on θ is brought about by the fact that, in the case in question, the coefficient j_ν includes the *stimulated* emission. The latter is determined by the radiation density in the direction concerned [formula (5.42)], and consequently depends on the angle θ. Taking account of the fact that only the fraction $\mathrm{d}\omega/4\pi$ of the total spontaneously emitted radiation is sent out in the beam $\mathrm{d}\omega$, we can now write, in accordance with (5.50) and (5.45),

$$j_\nu(\theta)\,\varrho\,\mathrm{d}\nu\,\mathrm{d}\omega = n_{r+1}\,\beta_\nu\,v\left(\frac{1}{4\pi} + \frac{c^3}{8\pi h\nu^3}\,\frac{I_\nu(\theta)}{c}\right) h\nu\,\mathrm{d}n_e\,\mathrm{d}\omega\,, \quad (5.57)$$

since a quantum $h\nu$ is emitted in every recombination.

Introducing, by means of Saha's formula (5.11), the quantity n_r in place of n_{r+1}, replacing n_r in its turn by $n_{r,k}$ with the aid of (5.54), and using (5.7), (5.8), (5.37), (5.52) and (5.55), we obtain

$$j_\nu(\theta)\,\varrho\,\mathrm{d}\nu\,\mathrm{d}\omega = n_{r,k}\,k'_\nu\,e^{-h\nu/kT}\left\{\frac{2h\nu^3}{c^2} + I_\nu(\theta)\right\}\mathrm{d}\omega\,\mathrm{d}\nu\,. \quad (5.58)$$

Now, replacing the first term in the braces by means of Planck's Law (3.11), we obtain instead of (5.58)

$$j_\nu(\theta)\,\varrho\,\mathrm{d}\nu\,\mathrm{d}\omega = n_{r,k}\,k'_\nu\left\{(1 - e^{-h\nu/kT})\,B_\nu + e^{-h\nu/kT}\,I_\nu(\theta)\right\}\mathrm{d}\omega\,\mathrm{d}\nu\,. \quad (5.59)$$

We now introduce this expression (5.59) into the equation of transfer (5.56), replacing the quantity $\varkappa'_\nu\,\varrho$ in the latter by the equal* quantity $k'_\nu\,n_{r,k}$, in accordance with (5.20) and (5.22):

$$\begin{aligned}\cos\theta\,\mathrm{d}I_\nu(\theta)/\mathrm{d}h &= n_{r,k}\,k'_\nu\,I_\nu(\theta)\,(1 - e^{-h\nu/kT}) - n_{r,k}\,k'_\nu\,B_\nu\,(1 - e^{-h\nu/kT})\\ &= n_{r,k}\,k'_\nu\,(1 - e^{-h\nu/kT})\,(I_\nu(\theta) - B_\nu)\,. \qquad (5.60)\end{aligned}$$

* The quantities $n_{r,k}$ and ϱ here relate only to atoms in the ionisation state r and the excitation state k.

We now introduce the following quantity:

$$k_\nu = (1 - e^{-h\nu/kT})\, k'_\nu \,. \tag{5.61}$$

The equation of transfer (5.60) then takes the form

$$\cos\theta\, \mathrm{d}I_\nu(\theta)/\mathrm{d}h = n_{r,\,k}\, k_\nu \,\{I_\nu(\theta) - B_\nu\}\,, \tag{5.62}$$

or, introducing again the absorption coefficient referred to unit mass,

$$\cos\theta\, \mathrm{d}I_\nu(\theta)/\mathrm{d}h = \varrho\, \varkappa_\nu \,\{I_\nu(\theta) - B_\nu\}\,. \tag{5.63}$$

We have thus obtained the original equation of transfer (3.12).

Consequently, in using the equation of transfer (3.12), the absorption coefficient which takes account only of the *stimulated absorption** must be modified in accordance with formula (5.61). It is quite clear also that the relation (5.61) is applicable to photo-ionisation from any level. Hence, to take the *stimulated emission* into account, we must multiply the whole expression (5.34) by $(1 - e^{-h\nu/kT})$.

We shall make two comments regarding the result just obtained. Firstly, it is easily shown that we obtain the same relation (5.61) by considering not processes of photo-ionisation, but *discrete* transitions between the different quantum orbits of the electron. Secondly, the use of the basic formula (5.61) presupposes the applicability of the formulae of Saha and Boltzmann and of Maxwell's distribution, i. e. the existence of thermodynamic equilibrium.

We now return to our main topic, that of the absorption coefficient. Using (5.61) and (5.34), we can write the following expression for the absorption coefficient referred to one $r+1$ times ionised atom:

$$k_\nu = \frac{C_0 Z^2 p_e}{T^{\frac{3}{2}}\, \nu^3} \left\{ \frac{2\, hRZ^2}{kT} \sum_{n_0}^{\infty} \frac{g'}{n^3}\, e^{hRZ^2/n^2kT} \right\} (1 - e^{-h\nu/kT})\,. \tag{5.64}$$

6. **Free-free transitions. Final results.** The absorption coefficient just written takes account only of *photo-ionisation processes* from different levels. These transitions are often called **bound-free**, and those converse to them, i. e. *recombinations*, are called **free-bound.** However, besides these transitions, there are others which are called **free-free** transitions. Classical electrodynamics and quantum mechanics tell us that a system consisting of an ion and an electron moving in a hyperbolic orbit in its field can both absorb and emit radiant energy. In the first case the energy of the system increases, and in the second it diminishes. The parameters of the hyperbolic trajectory of the electron are changed as a result of every absorption or emission.

---

* Corresponding to the coefficient B_{ik} in the case of discrete transitions.

The quantum theory gives for these processes the following expression for the absorption coefficient, again referred to one $r+1$ times ionised hydrogen-like atom:

$$k_\nu = \frac{C_0 Z^2 p_e}{T^{\frac{3}{2}} \nu^3} (1 - e^{-h\nu/kT}) \, g'' , \tag{5.65}$$

where the factor which takes account of the negative absorption process is already included (in the parentheses). In the expression (5.65), g'' is a correction coefficient, which in the majority of cases differs little from unity. It may be estimated from the formula

$$g'' = 1 + 0.1728 \left(\frac{\nu}{R Z^2}\right)^{\frac{1}{3}} \left[1 + \frac{2kT}{h\nu}\right]. \tag{5.66}$$

The absorption coefficient (5.65) corresponding to free-free transitions is generally of importance only for the far infra-red region of the spectrum. In the visible and photographic regions it is noticeable only at very high temperatures (usually above 50,000° to 100,000°).

According to the definition (2.13) of the absorption coefficient, the latter has the *additive property* [which we have used in formula (5.29)]. Hence, in order to take account of both bound-free and free-free transitions, we must add the two coefficients (5.64) and (5.65); as a result we obtain the following expression, which determines the absorption coefficient for hydrogen-like atoms:

$$k_\nu = \frac{C_0 Z^2 p_e}{\nu^3 T^{\frac{3}{2}}} \left\{ \frac{2hRZ^2}{kT} \sum_{n_0}^{\infty} \frac{g'}{n^3} e^{hRZ^2/n^2kT} + g'' \right\} (1 - e^{-h\nu/kT}) . \tag{5.67}$$

This absorption coefficient is referred to one $r+1$ times ionised atom. If we introduce the degree of ionisation x by means of the first of the relations (5.18), then, according to the general formula (5.23), the absorption coefficient referred to one hydrogen-like ion (taking account of all ionisation states) is

$$k_\nu = \frac{C_0 Z^2 x p_e}{\nu^3 T^{\frac{3}{2}}} \left\{ \frac{2hRZ^2}{kT} \sum_{n_0}^{\infty} \frac{g'}{n^3} e^{hRZ^2/n^2kT} + g'' \right\} (1 - e^{-h\nu/kT}) . \tag{5.68}$$

The formulae (5.67) and (5.68) cannot be used if the state of the matter deviates greatly from thermodynamic equilibrium. In this case we must start from the basic formula (5.24), in which the absorption coefficient does not depend on the external conditions.

As an illustration of formula (5.68) we give a diagram (Fig. 12a) which shows the variation of $\varkappa_\nu$ for hydrogen at $T = 10{,}370°$ and $p_e = 244$ bars. In this figure the absorption coefficient is that referred to unit mass; this is obtained, according to (5.22), by dividing (5.68)

by the mass of a hydrogen atom, it being of course remembered that for hydrogen $Z = 1$. The dependence of $\log_{10} \varkappa_\nu$ on wavelength is given for various temperatures in Fig. 12(b). The coefficient $\varkappa_\nu$ is here referred

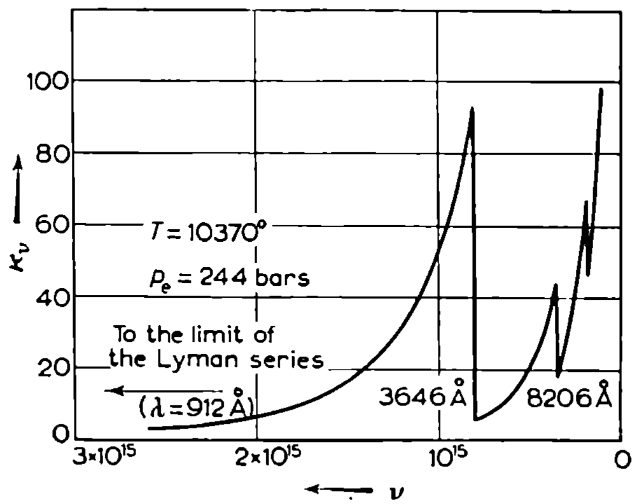

FIG. 12(a)

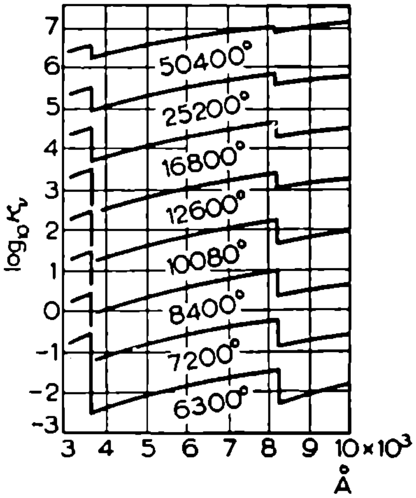

FIG. 12(b)

to one gram of neutral hydrogen. In order to obtain the absorption coefficient referred to 1 g of hydrogen (taking account of both the neutral and the ionised state), we must multiply the numbers in Fig. 12(b) by $(1 - x_\mathrm{H})$, where x_H is the degree of ionisation of hydrogen.

7. **Non-hydrogen-like atoms. Negative hydrogen ions. The scattering of radiation by free electrons.** All the results we have obtained, and, in particular, the original formula (5.24), are applicable to hydrogen-like atoms, i. e. to the atoms H, He II, Li III, Be IV, etc. The application to other elements of formula (5.24) and all those derived from it is unjustified. In these cases the absorption coefficient must be deduced either on the basis of experiment or from quantum-mechanical calculations. The deviations of the k_ν, so found, from the values obtained by using (5.24) are the smaller, the larger n, i. e. the greater the dimensions of the orbit from which the photo-ionisation takes place. Hence the application of (5.24) for levels with fairly high excitation potentials very often does not involve any considerable error. On the other hand, the application of this formula to photo-ionisation from the ground level (or from levels fairly close to it) may give results which differ by a factor of 10 or even 100 from the true values. Errors of the same order may result from the application of formula (5.24) to (non-hydrogen-like) atoms or ions which have only one electron outside the closed shells. Such systems, for example K I (neutral potassium), are called systems with one valency electron.

We shall here give a short survey of some data for the atoms which are most important from the astrophysical point of view. We shall

consider photo-ionisation from the ground level (absorption beyond the limit of the principal series).

The variation of k'_ν just beyond the limit of the principal series for helium is shown in Fig. 13. Here k'_ν decreases with increasing ν more slowly than ν^{-3}. The variation of k'_ν is shown in Fig. 14 for C, N, O and F. The energy of the ejected electrons in electron-volts is

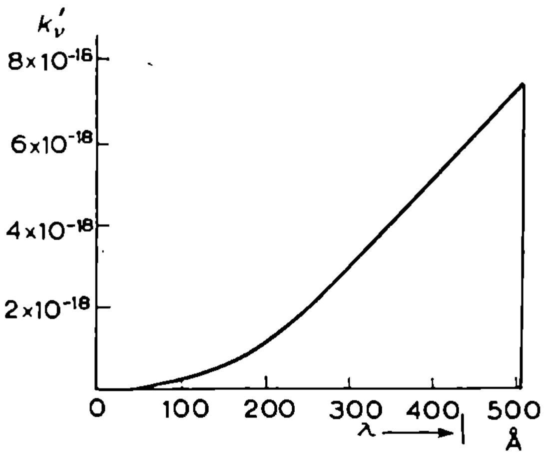

Fig. 13

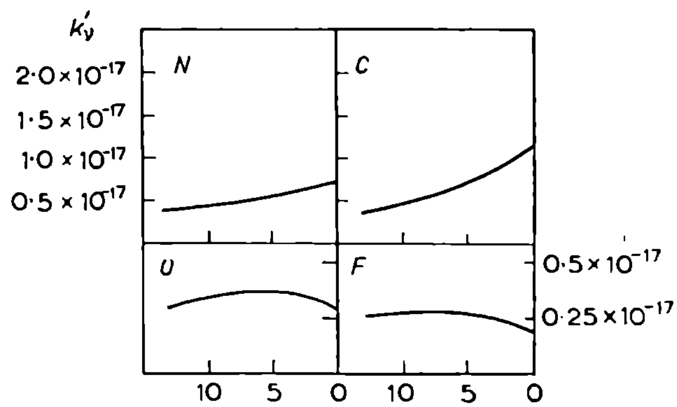

Fig. 14

placed on the axis of abscissae. For O and F the difference from the ν^{-3} law is particularly large, and we can even observe a kind of maximum.

For sodium, k_ν is $3{\cdot}1 \times 10^{-19}$ cm^2 just beyond the limit of the principal series. The further decrease of k_ν with frequency is more rapid than ν^{-3}. For calcium, k_ν follows closely the ν^{-3} law; just beyond the limit of the principal series, $k_\nu = 2{\cdot}5 \times 10^{-17}$ cm^2. For potassium, the existing calculations are still insufficiently accurate, but here k_ν is apparently considerably less than the value given by formula (5.24). For the remaining elements in Table 1 no calculations have yet been made. There are also hardly any calculations of k_ν for the photo-ionisation of the atoms mentioned from excited levels. However, in this case it is often possible to use formula (5.24), in view of the "hydrogen-likeness" of the orbits concerned.

There are almost no quantitative data on the values of k_ν for ions. The hydrogen-like ions (such as He II, Li III, etc.) form an exception, and so do those of a small number of selected elements.

As with hydrogen-like ions, all the other elements (neutral and ionised) are characterised by absorption consisting of separate bands superimposed on one another. In order to determine the absorption coefficient for frequency ν, we must take account of all these bands, though of course we need take into consideration only those bands whose edges lie in the direction of longer wavelengths from a given frequency. The absorption coefficient is constructed in the same way

as for hydrogen-like atoms. That is, the formulae of Boltzmann and Saha are used, and negative absorption is taken into account according to formula (5.61).

In 1939 it was discovered that, beginning from stars of class A 2 to F 0 and continuing to stars of later spectral classes, the absorption of radiation by *negative hydrogen ions* becomes more and more important, and finally is the dominant factor (for the Sun). A **negative hydrogen ion** is the name given to a system consisting of a neutral hydrogen atom with an additional electron attached to it. The absorption processes for this system are the same as for ordinary atoms. There are bound-free transitions, i. e. photoelectric absorption, and free-free transitions, i. e. absorption of radiation by the system consisting of a neutral hydrogen atom and a free electron moving near it in an open orbit. However, there is a certain difference between a negative hydrogen ion and ordinary atoms, consisting in the fact that in the case of the former there is practically only one important absorption band. Since *photoelectric absorption by the negative hydrogen ion* plays the chief part in the most important regions of the spectrum, we give separately in Fig. 15 (a) the absorption coefficient relating to this absorption only. It

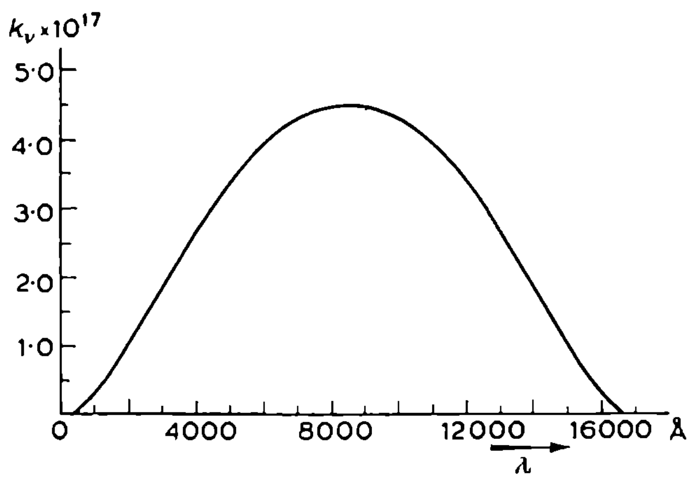

Fig. 15 (a)

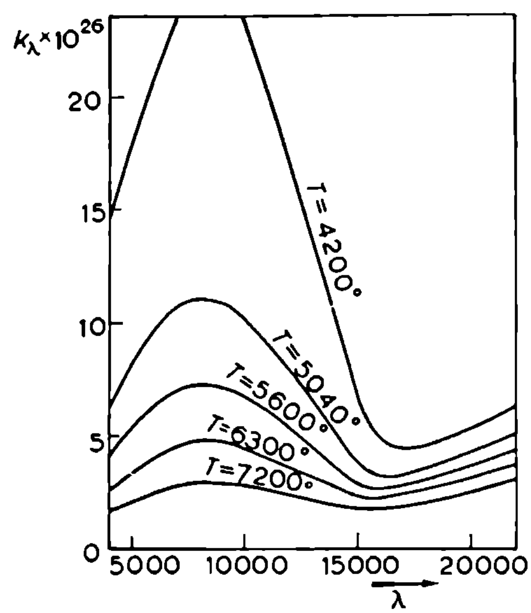

Fig. 15 (b)

is referred to one negative hydrogen ion. We notice that the ionisation potential for this system, which determines the edge of the absorption band, is 0·75 eV. The absorption coefficient of the negative hydrogen ion, taking account of free-free transitions, is shown in Fig. 15 (b). It has been constructed for a series of temperatures and is referred to one neutral hydrogen atom and unit electron pressure (1 bar). In order to find k_λ for any p_e, the k_λ found from the graph must be multiplied by this value of p_e.

There are no analytical expressions for the absorption coefficient of the negative hydrogen ion. The quantities k_ν for it are taken from the appropriate tables [31]. Besides negative hydrogen ions (denoted by H^-), there now exist calculations for negative oxygen ions.

In the photospheres of hot stars of classes O and B, where very few neutral hydrogen atoms remain, because of the high temperature, processes of *scattering of radiation by free electrons* begin to play an important part. However, it is necessary to bear in mind that we are here concerned not witht rue absorption but with scattering of radiation. The process of scattering amounts to the fact that the scattering particles simply change the direction of the quanta falling on them. The transformation of radiant energy into thermal energy and back does not occur here. Hence it is impossible in this case to apply directly Kirchhoff's Law (3.10) for the emission coefficient.

The scattering coefficient, like the absorption coefficient, is determined by the ordinary expression (5.23). If there are n_e free electrons in 1 cm^3 and the scattering coefficient referred to one free electron is s_e, then the attenuation of the intensity I_ν owing to scattering by free electrons is determined by the expression

$$\mathrm{d}I_\nu = -I_\nu\, s_e\, n_e\, \mathrm{d}h\ , \tag{5.69}$$

where

$$s_e = \frac{8\pi}{3}\left(\frac{e^2}{m_e c^2}\right)^2 . \tag{5.70}$$

The numerical value of s_e is $0{\cdot}665 \times 10^{-24}$ cm^2. It follows from (5.70) that the scattering by free electrons is *independent of the wavelength*. The scattering of radiation by protons, which are present in large numbers in the atmospheres of hot stars, is negligible, since the scattering coefficient for protons is considerably less than s_e.

When the absorption of radiation in stellar photospheres is determined not by one element but by two or more, it is necessary to introduce the *total absorption* in accordance with formula (5.23). If there are n_1 absorbing atoms of some element, with atomic absorption coefficient $k_\nu^{(1)}$, in 1 cm^3, n_2 atoms with absorption coefficient $k_\nu^{(2)}$, etc., then the total absorption in the frequency ν is determined, according to (5.23), by

$$\mathrm{d}I_\nu = \sum_s \mathrm{d}I_\nu^{(s)} = -I_\nu\, \mathrm{d}h \sum_s n_s\, k_\nu^{(s)} . \tag{5.71}$$

If we now wish to refer all absorption coefficients to one atom of the most abundant element, for instance hydrogen, formula (5.1) must be used. In this case, we find from (5.71)

$$\mathrm{d}I_\nu - -I_\nu\, n_1\, \mathrm{d}h \sum_s a_s\, k_\nu^{(s)} . \tag{5.72}$$

We shall call the quantity

$$k_\nu = \sum_s a_s k_\nu^{(s)} \tag{5.73}$$

the **total absorption coefficient.** According to formula (5.2) we could also have referred the total absorption coefficient to the total number of all atoms in 1 cm^3. However, since the number of hydrogen atoms n_{H} in 1 cm^3 considerably exceeds that of all atoms of other elements (in 1 cm^3), in practice $n_1 \approx n_{\mathrm{H}} \approx n$.

Having acquainted ourselves with the main outlines of the theory of absorption coefficients, we can now turn to the question of the energy distribution in the continuous spectra of stars.

Chapter 6. The distribution of energy in the continuous spectra of stars for an absorption coefficient depending on the frequency

1. **The absorption of radiation in the photospheres of stars of various spectral classes.** In Chapter 4 we came to the conclusion that the divergence between the theory there considered and the results of observation must be due mainly to the dependence of the absorption coefficient in stellar photospheres on the frequency. For this reason we have considered in Chapter 5 the most important facts concerning absorption coefficients. In order to facilitate the solution of further problems, we must also consider which atoms and ions are the main sources of absorption in the photospheres of stars of various spectral classes. In this we must be guided by formula (5.23), according to which the attenuation of the radiation is proportional to the number of absorbing atoms in 1 cm^3 and to the atomic absorption coefficient of these atoms. Starting from this, we shall now consider the part played by the various atoms and ions. We begin with stars of the spectral classes M and N. Direct spectroscopic observations show that practically the entire continuous spectrum of these stars is strongly distorted by bands due to various molecular compounds. Taking account of the absorbing effect of these bands involves very great difficulties. Moreover, the attenuation of the radiation by molecular bands is apparently due, in the majority of cases, to scattering, and not to true absorption.

Apart from molecular absorption, the absorption by negative hydrogen ions H^- may, apparently, play some part, though a small one, in the photospheres of the stars mentioned. The effectiveness of this source of absorption is the greater, the greater the number n_{H^-}, i. e.

the number of negative hydrogen ions in 1 cm^3, and this is greater the greater the number of neutral hydrogen atoms in 1 cm^3 and the greater the number of free electrons in 1 cm^3. The first condition is satisfied for all stars whose effective temperature T_e is less than 7000° to 8000°. In the atmospheres of these stars practically all the hydrogen is neutral. The second condition is not fulfilled in the photospheres of stars of the spectral classes M, N and R. A study of absorption lines shows (see Table 8, Chapter 15) that in the photospheres of stars of these classes the value of n_e is (relatively) small. Only in the atmospheres of K stars, where n_e is already fairly large (n_e increases with T_e, since the degree of ionisation of the atoms also increases), does the part played by negative hydrogen ions become important, while in the photosphere of the Sun (dG 3) it is the chief contribution to the absorption.

As the effective temperature increases further, the part played by H^- ions as a source of absorption begins to decrease again. This is chiefly caused by the increase in the absorbing power of atomic hydrogen H (see below). Moreover, as the temperature increases, the ionisation of H^- ions, i. e. their dissociation into neutral H atoms and electrons, increases. However, this factor is much less important, since the ionisation potential of the H^- ion is 0·75 eV, and with such a low ionisation potential the degree of ionisation changes comparatively little with T, according to Saha's formula (5.11). This change is slowest at fairly high temperatures, for at such temperatures the index of the exponential in Saha's formula is particularly small, and consequently even a large change in T has little effect. However, as T_e increases further, the ionisation of both H^- ions and of neutral hydrogen atoms begins to become important. Direct calculation shows that the part played by H^- ions becomes small in the photospheres of normal stars of class A 0, though in the photospheres of white dwarfs, where the electron pressure is very great, these ions are apparently important even for the class mentioned.

In conclusion, it should be mentioned that the electrons which, together with neutral hydrogen atoms, form H^- ions appear at the lower temperatures chiefly on account of the ionisation of elements whose ionisation potentials are relatively low. Among these elements are Na ($\chi_0 = 5{\cdot}12$ eV), Mg ($\chi_0 = 7{\cdot}61$ eV), Ca ($\chi_0 = 6{\cdot}09$ eV), etc. On the other hand, hydrogen, although it is quantitatively the predominating element, has a relatively high ionisation potential (13·59 eV). Hence its ionisation at low temperatures will give only a negligible number of free electrons. The same applies to He, N and O. Only at temperatures above 6000° or 7000°, with the electron pressures existing in stellar photospheres (see Table 8, Chapter 15), does hydrogen begin to become an important source of free electrons.

Let us return to the stars of classes M and N. Besides the molecular absorption and that by H^- ions, the absorption of radiation by atoms of metals must also play a certain part in these stars. However, no calculations have yet been made which could give a quantitative estimate of the relative importance of molecules, H^- ions and metals in the photospheres of stars of classes M and N.

Finally, we shall say a few words on the part played by hydrogen atoms. Despite the fact that practically all such atoms in the photospheres of late-type stars are neutral, the part they play is completely negligible. The absorption of radiation by neutral hydrogen atoms up to the limit of the Lyman series at $\lambda = 912$ Å is caused by photoionisation from excited levels. However, even for the first excited level with $n = 2$, the excitation potential of hydrogen atoms is very large, being 10·16 eV. Hence the number of neutral hydrogen atoms in excited states is completely negligible at low temperatures. Thus, for example, even for $T = 5700°$ we have seen that $n_2/n_1 = 4{\cdot}2 \times 10^{-9}$. For stars of classes M and N this ratio is several orders of magnitude smaller. Hence we can neglect, in this case, the part played by neutral hydrogen. It is true that we must not forget the existence of neutral hydrogen atoms in the ground state. By virtue of the large percentage content of hydrogen, these atoms will be very many in number. However, it follows from physical considerations, and from the analysis of the equations of the theory of radiative equilibrium, that the spectral regions which play the greatest part in establishing the law of variation of the temperature with depth are those in which the absorption coefficient and the intensity of the photospheric radiation are both fairly large. This is seen, in particular, from equation (3.14). In low-temperature stars, however, the intensity of radiation is particularly small in the spectral region where photo-ionisation of neutral hydrogen from the ground state takes place. In fact, the whole of this region lies in the far ultra-violet beyond the limit of the Lyman series at $\lambda = 912$ Å, i. e. in the **Lyman continuum.** Hence we may completely neglect, from this point of view, the part played by neutral hydrogen in the photospheres of low-temperature stars.

Let us now pass to hotter stars than those of classes M and N. We shall consider, in particular, stars of the solar type, where n_e is already quite large.

In the photospheres of stars of the solar type, molecular absorption no longer plays an important part. Direct calculations show that the absorption of radiation by excited metal atoms is also completely negligible in comparison with the absorption by H^- ions. Hence negative hydrogen ions are the main source of absorption in a very large range of wavelengths, including the infra-red, the visible and the near ultra-

violet (up to 4000 Å). The metals apparently play an important part only for wavelengths less than 3000 Å, i. e. in the region where they are photo-ionised from their ground levels. For wavelengths less than 4000 Å, the attenuation of the photospheric radiation by absorption lines, which are very strongly concentrated here, is also important. (It must be borne in mind that in many absorption lines the chief process is not true absorption, but scattering of radiation.) This also follows directly from observations, which show just beyond $\lambda = 4000$ Å an anomalously sharp decrease in the intensity of the continuous spectrum. The influence of absorption lines must also be very considerable in the far ultra-violet region of the spectrum up to the limit of the Lyman series, since in this region are concentrated a large number of strong resonance lines of various elements, i. e. lines caused by transitions from the lowest level. The lines of the Lyman series will be particularly strong. However, inasmuch as the intensity of radiation in the region considered is relatively small, the part played by the far ultra-violet regions of the spectrum in the energy balance of the solar photosphere cannot be especially large.

For the reasons just given, the absorption by neutral hydrogen atoms is unimportant in the solar photosphere. Thus the main source of absorption in the photospheres of solar-type and similar stars is formed by the H^- ions. The absorption by metals (photo-ionisation from the ground levels) and scattering of radiation in absorption lines are added in the ultra-violet region of the spectrum.

Let us now pass to still hotter stars. As the temperature increases, the number of excited neutral hydrogen atoms continually increases. When this number becomes comparable with the number of negative hydrogen ions*, both sources of absorption become equally effective. A comparison of Table 2 and Fig. 15(a) shows that in both cases the atomic absorption coefficient in the most important spectral regions is of the same order (10^{-17} cm^2).

When the temperature increases further, the number of excited hydrogen atoms begins to exceed the number of H^- ions, and neutral hydrogen becomes the substance determining the absorption. The absorption by H^- ions is apparently already inconsiderable in the photospheres of stars of class A 0. Consequently, a continuous transition from H^- ions to neutral hydrogen atoms takes place in the photospheres of stars of classes F and A.

As the temperature increases, the part played by absorption lines of metals in the ultra-violet decreases. It is apparently already small in stars of class A. In the photospheres of stars of classes A 0 to B 2

* The number of H^- ions n_{H^-} is always some orders of magnitude less than the total number of neutral H atoms (in all states), since n_{H^-} is limited by the number of free electrons.

the main absorbing element is *neutral hydrogen*. In the photospheres of giant stars of these classes, the scattering of radiation by *free electrons* begins to play a noticeable part.

In the photospheres of stars of classes O, B 0 to B 2, the absorption of radiation by *helium* atoms (in the ultra-violet region of the spectrum) is added to that by neutral hydrogen atoms. Moreover, in the photospheres of these stars the scattering of radiation by free electrons becomes an important factor; since hydrogen is almost completely ionised, there are very many electrons, and few neutral hydrogen atoms. However, it must be remembered that electron scattering is not a process of true absorption, and Kirchhoff's Law (3.10) cannot be applied to it.

It follows from the above discussion of absorption conditions that these conditions are mainly determined by the degree of ionisation of the atoms. The ionisation, in turn, is determined not only by the temperature, but also by the electron pressure. [As is seen from Saha's formula (5.11), the effect on x_r of a change in the temperature T is greater than the effect of the electron pressure p_e.] In general, it follows from (5.11) that an increase in T has the same effect as a decrease in p_e.

Let us now pass to the theory of radiative equilibrium with an absorption coefficient depending on the frequency. The very great mathematical complexity of the solution of this problem is increased by the fact that different kinds of absorption occur in the photospheres of stars of various classes, and the form of the dependence of $\varkappa_\nu$ on the frequency also varies. Hence the solution of this problem has to be carried out separately for each case. It is therefore necessary to introduce various simplifying assumptions or to use the method of successive approximations.

2. Radiative equilibrium in the photospheres of stars of classes A 0 to B 2. As an example, let us consider the radiative equilibrium in the photospheres of stars of classes A 0 to B 2. (We shall discuss stars of other spectral classes later.) In the case under consideration, the problem is simplified by the fact that the absorption of radiation is determined mainly by neutral hydrogen (see below regarding the scattering of radiation by electrons in these photospheres).

We shall solve the problem proposed by means of the method of successive approximations*, using the same process of averaging the intensity (over direction) which we discussed at the end of Chapter 3. (In Chapter 4 we gave arguments to show that this process is quite suitable for the case here considered.)

* The theory expounded below was constructed by É. R. Mustel' [94] — *Editor.*

For the coefficient of absorption by hydrogen atoms, we can write, using (5.68),

$$\varkappa_\nu = \Phi(p_e, T)\, \Psi(\nu, T)\,, \tag{6.1}$$

where

$$\Phi(p_e, T) = \frac{C_0\, x\, p_e}{m_{\mathrm{H}}\, T^{\frac{3}{2}}}\,, \tag{6.2}$$

$$\Psi(\nu, T) = \left[\frac{2hR}{kT} \sum_{n_0}^{\infty} \frac{g'}{n^3}\, e^{hR/n^2kT} + g''\right] \frac{1 - e^{-h\nu/kT}}{\nu^3}\,, \tag{6.3}$$

where in using (5.68) we have taken into account that for hydrogen $Z = 1$. In passing from k_ν to $\varkappa_\nu$ in accordance with (5.22) we have denoted the mass of the hydrogen atom by m_{H}.

In virtue of (6.1), the expression for the element of optical depth takes the form

$$\mathrm{d}\tau_\nu = \Phi(p_e, T)\, \Psi(\nu, T)\, \varrho\, \mathrm{d}h\,. \tag{6.4}$$

In order to simplify the analysis, we introduce a new variable ζ as follows:

$$\mathrm{d}\zeta = \tfrac{3}{2}\, \Phi(p_e, T)\, \varrho\, \mathrm{d}h\,. \tag{6.5}$$

We then have

$$\tfrac{3}{2}\, \mathrm{d}\tau_\nu = \Psi(\nu, T)\, \mathrm{d}\zeta\,, \tag{6.6}$$

$$\tfrac{3}{2}\, \tau_\nu = \int_0^\zeta \Psi(\nu, T)\, \mathrm{d}\zeta\,. \tag{6.7}$$

By means of (6.6) and (6.7) the equation (3.55) and the expressions (3.63) and (3.64) can be rewritten in the following form:

$$\frac{1}{\Psi(\nu, T)} \frac{\mathrm{d}J_\nu}{\mathrm{d}\zeta} = \tfrac{1}{2} H_\nu\,, \tag{6.8}$$

$$I_\nu = \exp\left[\int_0^\zeta \Psi(\nu, T)\, \mathrm{d}\zeta'\right] \int_\zeta^\infty \Delta_\nu \exp\left[-\int_0^{\zeta'} \Psi(\nu, T)\, \mathrm{d}\zeta''\right] \mathrm{d}\zeta' \tag{6.9}$$

and

$$I'_\nu = \exp\left[-\int_0^\zeta \Psi(\nu, T)\, \mathrm{d}\zeta'\right] \int_0^\zeta \Delta_\nu \exp\left[\int_0^{\zeta'} \Psi(\nu, T)\, \mathrm{d}\zeta''\right] \mathrm{d}\zeta'. \tag{6.10}$$

From (6.3) and (3.11), the function

$$\Delta_\nu = B_\nu\, \Psi(\nu, T) = \frac{2h}{c^2} \left[\frac{2hR}{kT} \sum_{n_0}^{\infty} \frac{g'}{n^3}\, e^{hR/n^2kT} + g''\right] e^{-h\nu/kT} \tag{6.11}$$

is a definite function of temperature and frequency. In order not to confuse the limit of integration with the independent variable of integration, we have added a prime to the latter.

It follows from the expressions (6.8) to (6.11) that the problem reduces to finding the dependence of T on ζ. In the case considered, the variable ζ completely replaces the variable τ which we used in Chapter 4.

Using the equation

$$\frac{dJ_\nu}{d\zeta} = \frac{dJ_\nu}{dT} \frac{dT}{d\zeta} \tag{6.12}$$

and integrating (6.8) over the whole spectrum, we obtain the basic differential equation of our problem:

$$\frac{dT}{d\zeta} = \frac{H}{2\int\limits_0^\infty \frac{1}{\Psi(\nu, T)} \frac{dJ_\nu}{dT}\, d\nu} \tag{6.13}$$

Since the derivative of the function J_ν with respect to temperature is unknown, the equation (6.13) has to be solved by the method of successive approximations. As a first approximation we use the expression for the derivative dJ_ν/dT which holds in sufficiently deep layers of the stellar atmosphere. We shall show, in fact, that as $\tau_\nu \to \infty$ the following equation holds:

$$dJ_\nu/dT = dB_\nu/dT\ . \tag{6.14}$$

We may remark in this connection that the equation $J_\nu = B_\nu$ and the equation (6.14) are satisfied in the presence of strict thermodynamic equilibrium, when the intensity of radiation is the same in all directions and is determined by Planck's formula (3.11). In this case the equation $J_\nu = B_\nu$ follows from formula (3.15). Thus the proof of the correctness of (6.14) is also a proof that strict thermodynamic equilibrium holds in the deep layers of the stellar atmosphere to a sufficient degree of accuracy.

To prove (6.14), we divide the equation (3.50) by I_ν and proceed to the limit as $\tau_\nu \to \infty$. We then obtain

$$\lim_{\tau_\nu \to \infty} \frac{H_\nu}{I_\nu} = 1 - \lim_{\tau_\nu \to \infty} \frac{I_\nu'}{I_\nu}\ . \tag{6.15}$$

We shall show that the left-hand side of (6.15) tends to zero as $\tau_\nu \to \infty$. To do so we recall that the integrated flux πH is constant at all depths. Since, according to (2.6), πH is the integral of πH_ν over ν, H_ν cannot increase without limit for any frequency*. As we penetrate into the

* It must be borne in mind that πH_ν is positive for all frequencies, and consequently it is not possible for one spectral region with a large positive H_ν to be compensated by another with a large negative H_ν.

stellar photosphere, only a transformation of the flux in the direction of greater frequencies takes place, and the quantities πH_ν themselves cannot exceed a certain limit (see also Fig. 18 and the explanation thereof). But since for $\tau_\nu \to \infty$ the temperature of the medium continuously increases, and therefore the intensity I_ν increases without limit, the limit of the ratio H_ν/I_ν is zero. Consequently, we can write

$$\lim_{\tau_\nu \to \infty} \frac{I_\nu'}{I_\nu} = 1 . \tag{6.16}$$

Let us now consider some point in the stellar photosphere at optical depth τ_ν, and measure the optical thickness upwards and downwards from this point (and not from the boundary of the photosphere). Then for $t_\nu \geq \tau_\nu$ and $t_\nu \leq \tau_\nu$, respectively

$$t_\nu^* = t_\nu - \tau_\nu,\ t_\nu^* = \tau_\nu - t_\nu . \tag{6.17}$$

The equations (3.63) and (3.64) can therefore be rewritten as

$$\left.\begin{aligned} I_\nu &= \int_0^\infty B_\nu\, e^{-\frac{3}{2} t_\nu^*}\ \tfrac{3}{2}\,\mathrm{d}t_\nu^* , \\ I_\nu' &= \int_0^{\tau_\nu \to \infty} B_\nu\, e^{-\frac{3}{2} t_\nu^*}\ \tfrac{3}{2}\,\mathrm{d}t_\nu^* . \end{aligned}\right\} \tag{6.18}$$

These expressions [taking account of what was said concerning formula (4.27)] enable us to regard I_ν and I_ν' as weighted means of B_ν.

Let now the temperature at the point selected be T. Then the whole of the photospheric layers lying above will have a lower temperature than T, and those lying below (i. e. towards the centre of the star) a higher temperature. The same inequality will be true of the functions B_ν themselves. Then we can write for the given point, by virtue of what has just been said about the intensities I_ν and I_ν',

$$I_\nu > B_\nu(T) > I_\nu' . \tag{6.19}$$

This inequality, together with (6.16), enables us to write

$$\lim_{\tau_\nu \to \infty} \frac{I_\nu'}{B_\nu} = 1 ,\quad \lim_{\tau_\nu \to \infty} \frac{I_\nu}{B_\nu} = 1 . \tag{6.20}$$

We now divide (3.51) by B_ν and pass to the limit as $\tau_\nu \to \infty$:

$$\lim_{\tau_\nu \to \infty} \frac{J_\nu}{B_\nu} = \tfrac{1}{2} \left\{ \lim_{\tau_\nu \to \infty} \frac{I_\nu}{B_\nu} + \lim_{\tau_\nu \to \infty} \frac{I_\nu'}{B_\nu} \right\} . \tag{6.21}$$

Introducing the expressions (6.20) into (6.21), we find that in fact, for $\tau_\nu \to \infty$, J_ν tends to B_ν, and this enables us to write (6.14).

Now, introducing (6.14) into (6.13), we obtain the equation for the first approximation:

$$\mathrm{d}T/\mathrm{d}\zeta = H/2\,Z(T)\,, \tag{6.22}$$

where

$$Z(T) = \int_0^\infty \frac{1}{\Psi(\nu, T)} \frac{\mathrm{d}B_\nu}{\mathrm{d}T}\,\mathrm{d}\nu \tag{6.23}$$

is a definite function depending only on the temperature T. Introducing the function $\Psi(\nu, T)$ from (6.3) and the derivative of B_ν with respect to the temperature in the integral in (6.23), and effecting the integration by numerical methods [95], we can determine $Z(T)$ for various T.

Having calculated the function $Z(T)$, we can integrate the equation (6.22) for a given H, i. e. for a given T_e. Since for $\zeta = 0$ the temperature T is equal to the boundary temperature T_0, this gives

$$\int_{T_0}^{T} Z(T)\,\mathrm{d}T = \tfrac{1}{2}\,H\zeta\,. \tag{6.24}$$

We must now determine the temperature T_0, which is a constant of integration. If we found the exact solution of the problem, the distribution of temperature obtained would give us the same flux πH, equal to σT_e^4, at all depths, so that the deviation from σT_e^4 is a measure of the incorrectness of the solution obtained. Hence we determine the constant of integration so that the solution found in the first approximation (i. e. the dependence of T on ζ) gives the prescribed flux σT_e^4 at the boundary, where $\zeta = 0$. According to (6.9), the intensity I_ν at the boundary of the photosphere is

$$(I_\nu)_0 = \int_0^\infty \Delta_\nu(T) \exp\,[-\int_0^\zeta \Psi(\nu, T)\,\mathrm{d}\zeta]\,\mathrm{d}\zeta\,. \tag{6.25}$$

At the boundary of the photosphere, where $I'_\nu \equiv 0$, we have, in accordance with (3.50) and (6.25),

$$\begin{aligned}(\pi H)_0 &= \pi \int_0^\infty (H_\nu)_0\,\mathrm{d}\nu = \pi \int_0^\infty (I_\nu)_0\,\mathrm{d}\nu \\ &= \pi \int_0^\infty \Big\{\int_0^\infty \Delta_\nu(T) \exp\,[-\int_0^\zeta \Psi(\nu,T)\,\mathrm{d}\zeta]\,\mathrm{d}\zeta\Big\}\,\mathrm{d}\nu.\end{aligned} \tag{6.26}$$

Let us assume that we have prescribed some surface temperature T_0. Then the temperature distribution can be found from the equation (6.24). Knowing T for every value of ζ, we can then calculate by means of numerical integration, in accordance with (6.3), (6.11) and (6.25), the quantities $(I_\nu)_0$ for various frequencies, and then by formula (6.26) the total flux $\pi(H)_0$, which also has to be calculated by numerical integration.

The quantity T_0 itself which is such that it allows the basic condition

$$(\pi H)_0 = \sigma\, T_e^4 \tag{6.27}$$

to be satisfied is found as follows. We take a series of three or four values for T_0. Let these be $T_0^{(1)}$, $T_0^{(2)}$, $T_0^{(3)}$, $T_0^{(4)}$, lying close to the expected value; for each of these values we integrate equation (6.24). Then, for each of the temperature distributions thus obtained, we find the value of $(I_\nu)_0$, and then we calculate, by means of (6.26), $(\pi H)_0$, i. e. the fluxes $(\pi H^{(1)})_0$, $(\pi H^{(2)})_0$, $(\pi H^{(3)})_0$ and $(\pi H^{(4)})_0$, corresponding to the selected values of T_0. We construct a graph, placing on the axis of abscissae $T_0^{(i)}$, and on the axis of ordinates the corresponding fluxes $(\pi H^{(i)})_0$ $(i = 1, 2, 3, 4)$. On joining the points obtained by an interpolation curve, we find on this graph the temperature T_0 which corresponds to the given flux $\sigma\, T_e^4$. This is the required surface temperature (Fig. 16). If, using this temperature, we again integrate equation (6.24) and calculate as before the quantities $\pi(I_\nu)_0$ and $(\pi H)_0$, then we should obtain for $(\pi H)_0$ the value $\sigma\, T_e^4$ (this serves as a check on the correctness of the T_0 obtained).

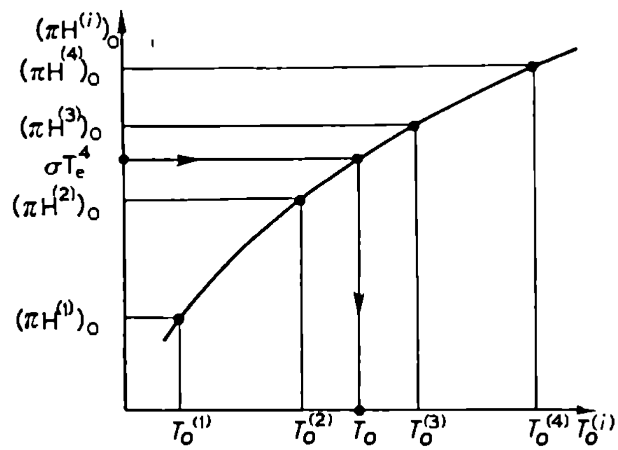

FIG. 16

The quantities $(I_\nu)_0$ for the T_0 finally obtained are to be identified with the distribution of energy in the spectrum of the total stellar radiation. For, according to (4.47), the mean intensity of the stellar radiation is $(H_\nu)_0$, and so, in virtue of (3.50), it is equal to $(I_\nu)_0$, since $I'_\nu \equiv 0$ at the boundary. Here, of course, we are speaking as yet of the first approximation.

The temperature distribution which we have obtained in the first approximation gives the prescribed flux $\sigma\, T_e^4$ at the surface of the star. However, a more detailed consideration of the question shows that, in finding T_0 in the manner described, we achieve the constancy of πH not only at the boundary itself (where $\zeta \equiv 0$), but also at great depths, with $\tau_\nu \to \infty$. Actually, from (6.14), the temperature gradient for these τ_ν which we find by means of (6.22) is practically exact. Moreover, direct calculations, as well as physical analysis, show that the effect of an error in the temperature of the outer layers on the temperature distribution in the deep layers is small. Then the constancy of the flux πH

and its equality to σT_e^4 are fulfilled in the first approximation both in the layers nearest the surface of the star and in the deep layers.

We now go to the second approximation. On the basis of the temperature distribution, i. e. the function $T(\zeta)$, found in the first approximation, we can determine the quantities I_ν and I'_ν from formulae (6.9) and (6.10) for a series of values of ζ, and therefore of T. Consequently we can determine J_ν from formula (3.51) for these values of ζ and T, and thence also $\mathrm{d}J_\nu/\mathrm{d}T$ by numerical differentiation. Thus we can introduce into equation (6.13) for each ζ not the quantities $\mathrm{d}B_\nu/\mathrm{d}T$, but the more exact quantities $\mathrm{d}J_\nu/\mathrm{d}T$, derived from the first approximation. The problem reduces to the numerical integration of equation (6.13), where $\mathrm{d}J_\nu/\mathrm{d}T$ is a function of ζ, and $\Psi(\nu, T)$ is again a function of the temperature distribution obtained. The surface temperature in the second approximation must be found in the same way as in the first approximation.

If necessary, we could also introduce the third approximation, and so on. However, actual calculations show that even the first approximation gives completely satisfactory results, i. e. a practically exact flux at all levels. Hence we can restrict ourselves to the first approximation.

3. **Comparison of theory and observation.** Numerical calculations have been carried out [95] in the first approximation, in accordance with the theory just explained, for effective temperatures $T_e = 10{,}500°$, 15,000° and 20,000°. We shall give here those results which are the most important from the theoretical standpoint.

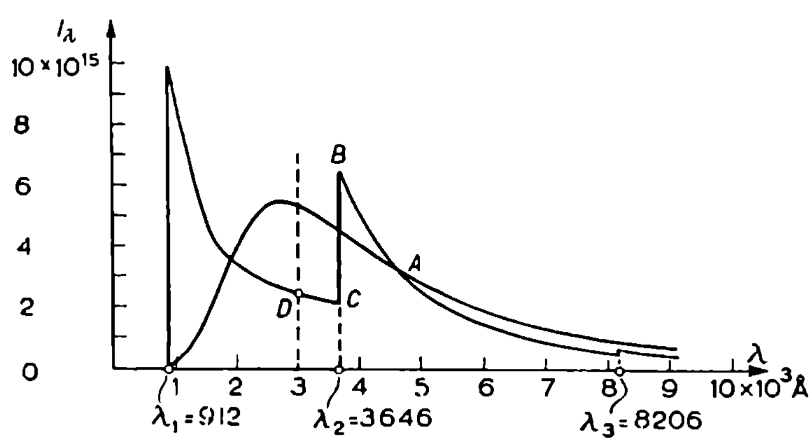

Fig. 17

Fig. 17 shows the distribution of energy, calculated in the manner described, in the continuous spectrum of a star with $T_e = 10{,}500°$ and also, for the sake of comparison, Planck's function B_λ for the same temperature of 10,500°.

The most characteristic property of stellar radiation is the presence of discontinuities in the mean intensity $\bar{I}_\lambda$ at the series limits. The origin of these discontinuities is easily explained. To do so, we return to Fig. 12. According to this figure, the absorption coefficient becomes very large just beyond the limit of the Balmer series. This means that, in the spectral region in question, we observe radiation coming from the uppermost layers of the stellar photosphere, where the temperature is relatively low (being close to the surface temperature of the star). Consequently, in agreement with Fig. 17 (segment CD), the intensity $\bar{I}_\lambda$ beyond the limit of the Balmer series will be relatively small. On this segment (close to the point C) the mean intensity of radiation $\bar{I}_\lambda$ is in fact close to the Planck intensity B_λ for the *surface temperature*. On the other hand, up to the limit of the Balmer series the absorption coefficient is relatively small. This means that here the radiation comes to us from fairly deep, relatively hot, layers of the stellar photosphere, and hence $\bar{I}_\lambda$ will be relatively high (segment AB in Fig. 17). The discontinuities at the limits of the other series are explained in the same way.

A second property of the theoretical curve in Fig. 17 is that the colour temperature T_c in some parts of the spectrum differs very greatly from the effective temperature, and in general varies with the wavelength. Thus, in the case considered, i. e. for $T_e = 10{,}500°$, in the region beyond the limit of the Balmer series, at $\lambda \approx 3350$ Å, the value of T_c is of the order of 11,000°, while in the region up to this series limit, at $\lambda \approx 4400$ Å, $T_c \approx 19{,}000°$, and finally, at $\lambda \approx 5500$ Å, the value of T_c is close to 15,000°. In other words, for stars with $T_e = 10{,}500°$ the form of $\bar{I}_\lambda$ is very different from the Planckian with $T = T_e$. This situation arises from two factors: firstly, from the temperature variation in the stellar photosphere, and secondly, from the dependence of $\varkappa_\nu$ on the frequency. For instance, if the temperature T inside the photosphere were constant and equal to T_e, then the function B_ν would be constant, and according to (4.49) we should have $\bar{I}_\nu = B_\nu(T_e)$ for every frequency, regardless of the dependence of $\varkappa_\nu$ on ν. [When the quantity B_ν is removed from the integrand in (4.49), the integral becomes equal to unity.] A dependence similar to that shown in Fig. 17 holds also for $T_e = 15{,}000°$ and even for $T_e = 20{,}000°$, the only difference being that, when T_e increases, the discontinuity at the limit of the Balmer series, defined by

$$D = \log_{10}(I_B/I_C)\,, \tag{6.28}$$

where I_B and I_C are the intensities at the points B and C respectively (see Fig. 17), decreases. Its value is shown in Table 3.

Table 3

T_e	D	T_0	$T_0^{(g)}$	Spectral class
10,500°	0·49	8430°	8830°	A 0
15,000°	0·22	11,880°	12,610°	B 5
20,000°	0·10	16,300°	16,820°	B 2

Fig. 18 shows the monochromatic flux H_ν for a star with $T_e = 10{,}500°$ and various ζ. The transformation of the flux H_ν in the direction of greater frequencies as we penetrate deeper into the photosphere, while the integrated flux remains constant, is clearly seen from this figure; we have mentioned this effect previously.

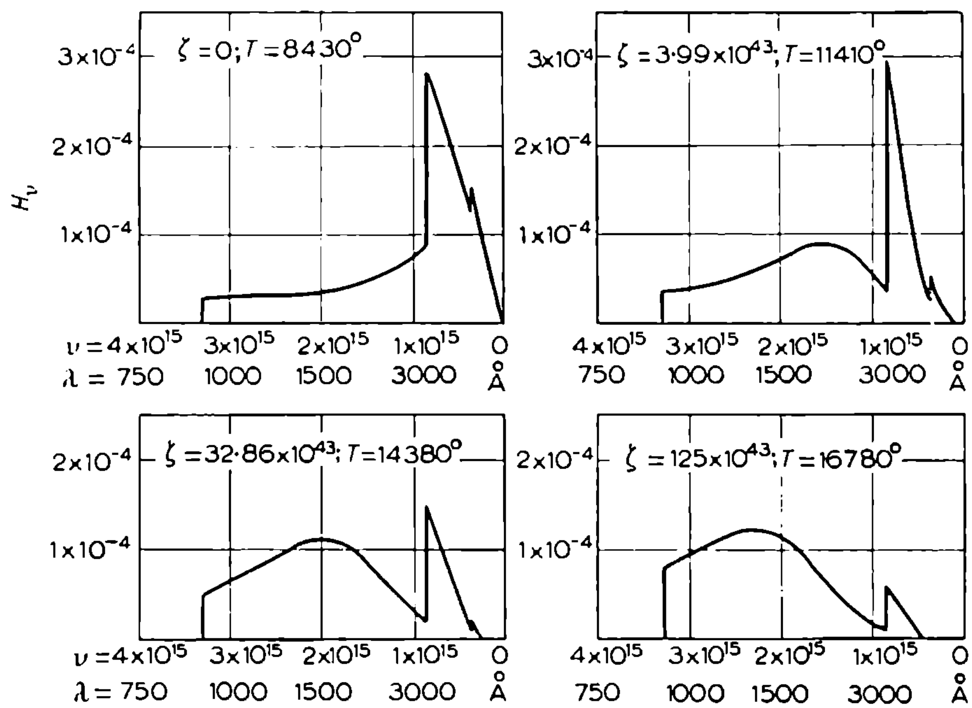

FIG. 18

A number of considerations (which we shall discuss in Part II), which are partly based on the results of the calculations explained above, give reason to suppose that effective temperatures $T_e = 10{,}500°$, 15,000° and 20,000° are to be ascribed to stars of the spectral classes A 0, B 5 and B 2 respectively. In this case the observations are in satisfactory agreement with the theoretical curves which give the variation of I_λ, both as regards the value of D, and as regards the values of T_c for the various spectral regions [168, § 52]. These facts indicate that our original suppositions about the existence of radiative equilibrium in the photospheres of stars of classes A 0 to B 2 are essentially correct. Our original assumption that the absorption of radiation is determined mainly by hydrogen is shown at the same time to be correct also.

In conclusion we may remark that, for the stars considered, the surface temperatures T_0 are lower than in the case of grey material. The surface temperatures calculated from formula (4.24), denoted by $T_0^{(g)}$, are given in the fourth column of Table 3. The inequality $T_0 < T_0^{(g)}$ is due to the fact that the surface temperature of a star is determined mainly by the regions of the spectrum where the intensity of radiation and the absorption coefficient are both large. In our case these are the frequencies lying just beyond the limit of the Balmer series. At the same time, we see from Fig. 17 that for these frequencies (segment CD) the inequality $\bar{I}_\lambda < B_\lambda(T_e)$ holds. This leads to the inequality $T_0 < T_0^{(g)}$. In general, the true values of T_0 may be even lower than those given in Table 3. In fact, we have seen that, in the case of grey material, the exact allowance for the variation of the intensity of radiation with direction diminishes T_0 for a given T_e. The same can happen when $\varkappa_\nu$ varies with frequency.

However, a number of considerations show that, for the stars we are considering, the use of the *mean* intensities I_ν and I'_ν cannot involve noticeable errors. In order to prove this, let us discuss the boundary of the photosphere, where these errors should be greatest.

The analysis of the equation of radiative equilibrium (3.16), as well as (6.13), shows that the depth distribution of temperature is determined mainly by those parts of the spectrum where the absorption coefficient and the mean intensity of radiation are large *at the same time* (see also Fig. 20 and the text relating to it). These parts of the spectrum lie *immediately* beyond the limits of the corresponding series (see, for example, the segment CD in Fig. 17). Moreover, for these frequencies (see Chapter 8), the intensity of radiation at the boundary, calculated by formula (3.38), is very close, for all $\theta \leq \frac{1}{2}\pi$, to the value $B_\nu(T_0)$, i.e. varies only slightly with θ. Thus, in using various methods of averaging the intensity over direction, we do not commit any important error.

On the other hand, for *grey* material, *all* regions of the spectrum (with a sufficiently high intensity, of course) play approximately the same part, and the intensity $I_\nu(\theta)$ at the boundary varies markedly with the angle θ. In this case the errors involved in the use of the mean intensities I_ν and I'_ν may in some cases be noticeable, even if not very large, as follows from Fig. 8.

The fact that the theory given above for pure hydrogen absorption (for various T_e) is fairly exact is seen as follows. As has already been said, when we have obtained the *final* temperature distribution in the photosphere of the star, we again find the flux $(\pi H_\nu)_0 = (\pi I_\nu)_0$ for various frequencies from formula (6.25), i. e. using the mean intensities, and the integration of the flux over the whole spectrum should give us $\pi H = \sigma T_e^4$.

On the other hand, knowing the temperature distribution, we can find the flux at the boundary, using the *exact* formula (3.38) for the intensity. In this case, having determined $I_\nu(\theta, 0)$ for various θ, we find the flux $(\pi H_\nu)_0$ from (2.9), where, of course, the integration is from $\theta = 0$ to $\theta = \frac{1}{2}\pi$, and not to $\theta = \pi$. Such calculations, which have been carried out for a photosphere with $T_e = 10{,}500°$, show that the values of $(\pi H_\nu)_0$ in the two cases are *very close* to each other for very different parts of the spectrum and, in particular, up to and immediately beyond the limit of the Balmer series. The use of the fundamental condition $\pi H = \sigma T_e^4$ is therefore sufficiently exact in the theory discussed above.

4. **Stars of spectral classes close to that of the Sun.** Let us now pass on to the consideration of stars of other spectral classes. Here it is necessary, as we have already said, to work out the particular method which is the simplest for each case separately. One of the most widely used methods is the introduction of some mean (over the spectrum) of the absorption coefficient $\varkappa_\nu$. This method is at its most effective when the ratio of $\varkappa_\nu$ to the mean absorption coefficient $\bar{\varkappa}$ introduced, i. e. the quantity

$$p_\nu = \varkappa_\nu / \bar{\varkappa} \,, \tag{6.29}$$

does not vary very greatly with frequency. This case corresponds fairly closely to the photospheres of stars of the solar type, where the absorption of radiation is mainly determined by negative hydrogen ions and where the change of $\varkappa_\nu$ in the wavelength range $\lambda = 3000$ Å to 7000 Å is relatively small (see Fig. 15) compared with the change of $\varkappa_\nu$ in the case of absorption by hydrogen (see Fig. 12). (The effective temperature of the Sun is 5710°.)

In the photospheres of stars of the solar type (or close to it), the quantity p_ν can depend only slightly on the depth. For, with a slight dependence of $\varkappa_\nu$ on ν, the ratio of $\varkappa_\nu$ to $\bar{\varkappa}$, for any method of averaging $\varkappa_\nu$, cannot differ markedly from unity, and so this ratio cannot change markedly with depth. By means of a fairly complicated analysis, S. Chandrasekhar has shown that, in the case where p_ν is completely independent of the depth, we can use the temperature distribution deduced for grey material, i. e., for example, the expression (4.23) or (4.26). The optical depth τ which appears in this expression must be determined by the mean absorption coefficient $\bar{\varkappa}$. The mode of formation of $\bar{\varkappa}$ from the values of $\varkappa_\nu$ for a given level is to be the following:

$$\bar{\varkappa} = \int_0^\infty H_\nu^{(1)} \varkappa_\nu \, \mathrm{d}\nu / H \,, \tag{6.30}$$

where $\pi H_\nu^{(1)}$ is the flux calculated for this level with an absorption coefficient independent of the frequency; the calculations are performed

for a prescribed $\pi H = \sigma T_e^4$. In other words, the solution (4.23) or (4.26) is again used to calculate $H_\nu^{(1)}$.

The energy distribution in the continuous spectrum of the emergent radiation, in the case considered, is found as follows. Because p_ν is constant, we can write for the optical depth

$$\tau_\nu = \int \varkappa_\nu \varrho \, dh = \int \varkappa_\nu \frac{\bar{\varkappa}}{\bar{\varkappa}} \varrho \, dh = p_\nu \int \bar{\varkappa} \varrho \, dh = p_\nu \tau . \tag{6.31}$$

The intensity of the emergent radiation at the boundary of the star is, according to (3.38),

$$I_\nu(0,0) = \int_0^\infty B_\nu \, e^{-p_\nu \tau \sec \theta} \sec \theta \, p_\nu \, d\tau . \tag{6.32}$$

The mean intensity of the observed stellar radiation, which according to (4.47) is equal to H_ν at the boundary, is

$$\bar{I}_\nu = (H_\nu)_0 = 2 \int_0^{\frac{1}{2}\pi} I_\nu(0,0) \cos \theta \sin \theta \, d\theta$$

$$= 2 \int_0^\infty B_\nu \, d\tau \int_0^{\frac{1}{2}\pi} e^{-p_\nu \tau \sec \theta} p_\nu \sin \theta \, d\theta . \tag{6.33}$$

Putting here $\sec \theta = w$, we obtain

$$\bar{I}_\nu = 2 \int_0^\infty B_\nu(T) \, E_2(p_\nu \tau) \, p_\nu \, d\tau , \tag{6.34}$$

where $E_2(x)$ is a particular case of the function $E_n(x)$, which is defined by the general relation

$$E_n(x) = \int_1^\infty \frac{e^{-xw}}{w^n} \, dw = x^{n-1} \int_x^\infty \frac{e^{-u} \, du}{u^n} . \tag{6.35}$$

The integration by parts of (6.35) leads to the following recurrence formula:

$$(n-1) \, E_n(x) = e^{-x} - x E_{n-1}(x) , \tag{6.36}$$

which gives in particular

$$E_2(x) = e^{-x} - x E_1(x) . \tag{6.37}$$

The quantity $E_1(x)$ is found in several tables (see, for instance, [63, p. 6]). Consequently, knowing $p_\nu \, \tau$, we can also determine $E_2(p_\nu \tau)$ from formula (6.37).

Thus, in order to find the intensity of the continuous spectrum from formula (6.34), we must know the quantities p_ν for a number of frequencies. The connection between T and τ is given, as we have already said, by means of the solutions (4.23) or (4.26), or by means of the solution

$$T^4 = \frac{1}{2} T_e^4 \left(\frac{1}{2}\frac{3}{2} + \frac{3}{2}\tau\right), \tag{6.38}$$

which combines the gradient of the solution (4.23) with the relation (4.25). The solution (6.38) is more exact than (4.23), but less exact than (4.26).

The method described has been applied by S. CHANDRASEKHAR and G. MÜNCH to the study of the Sun, for which $T_e = 5710°$. If the quantity p_ν for different depths were, as we suppose, in fact constant, its determination could be referred to any level. However, the supposition about the constancy of p_ν is of course somewhat schematic, since there is in reality a definite change of p_ν with depth. Hence the quantity p_ν must be determined for the depth which is most effective in forming the continuous spectrum of the star. This level in general corresponds to the photospheric layers where the local temperature T equals the effective temperature T_e (τ approximately 0·6). Calculations based on all these considerations, and carried out on the supposition that negative hydrogen ions are the only source of absorption in the solar photosphere, have led to the following results. It appears that theory and observation may be brought into satisfactory agreement for a fairly large wavelength interval, from $\lambda \approx 4000$ Å to 16,000 Å, if the theoretical value of $\bar{\varkappa}$, found by means of (6.30), is increased by a factor of 1.4. This latter circumstance is not unexpected, since it must be borne in mind that $\bar{\varkappa}$ is to be determined by averaging $\varkappa_\nu$ over the whole spectrum. At the same time we have already said that, in the ultra-violet region of the spectrum beyond 3000 Å, the absorption of solar radiation by metals owing to photo-ionisation from the ground levels must play an important part. It is clear that the introduction of absorption by metals must increase $\bar{\varkappa}$ in comparison with $\bar{\varkappa}(H^-)$.

There exists another method (see Section 7.3) which makes it possible to determine, from the observed dependence of $I_\nu(\theta,0)$ on the angle θ, the change of $\varkappa_\nu$ with frequency for various levels in the photosphere, i. e. for various local temperatures. This method, too, gives curves for $\varkappa_\nu$ which coincide very closely in the interval from 4000 to 16,000 Å with the curves shown on the right of Fig. 15. Thus negative hydrogen ions are without doubt the main source of absorption in the solar photosphere in this wavelength interval. Further, the agreement between theory and observation shows that the application of the theory of radiative equilibrium to the solar photosphere also is sufficiently justified. The divergence between theory and observation beyond $\lambda = 4000$ Å is

apparently due to the concentration of absorption lines, as we have already said. The position regarding the infra-red region with $\lambda > 16{,}000$ Å is still uncertain, since the measurements themselves of the energy distribution in this region are still very unreliable.

5. **Stars of other spectral classes. Stars with extended photospheres.** Since the theory of radiative equilibrium with the absorption coefficient $\varkappa_\nu(\mathrm{H}^-)$ satisfactorily explains the solar spectrum, it may be thought that the same will also be true for stars whose spectral class is not much different from that of the Sun (dG 3). However, for stars of late types, just as for stars of spectral classes earlier than that of the Sun, the application of the theory meets with various difficulties. In the former case these difficulties are due to molecular bands, which appear at the spectral class K 5 and reach a very great intensity in classes M, R, N and S. There is as yet no theory which takes account of the influence of these bands on the temperature distribution.

Another case is the *transition* between stars of the solar type and stars of the spectral classes F and A. Here the main difficulty is that in the photospheres of these stars both *negative hydrogen ions* and *neutral hydrogen* play an important part. Hence the value of p_ν is now a function of the depth. On the other hand, the representation of $\varkappa_\nu$ by the formula (6.1) is also incorrect in this case. As a result of all these difficulties there are as yet no calculations which are at all reliable.

In conclusion, let us consider the very hot O stars, and B stars of the early sub-classes. Here, as we have already said in the introduction to the present chapter, *scattering of radiation by free electrons* begins to play an important part, which increases continuously with temperature, as may be seen from the following considerations. Let 1 cm^3 contain n hydrogen atoms. Since the latter are largely ionised at the temperatures in question, hydrogen will be the chief source of free electrons, because of the high proportion of it in the atmosphere. Consequently, if n_p is the number of protons in 1 cm^3, we have

$$n_e \approx n_p = x n \,. \tag{6.39}$$

At fairly high temperatures, electrons coming from the ionisation of helium also play some part.

In accordance with (6.39) and (5.69), in the case of scattering by free electrons we have

$$\begin{aligned} \mathrm{d}I_\nu &= - I_\nu\, n_e\, s_e\, \mathrm{d}h = - I_\nu\, x\, n\, m_{\mathrm{H}}\, (s_e/m_{\mathrm{H}})\, \mathrm{d}h \\ &= - I_\nu\, x\, (s_e/m_{\mathrm{H}})\, \varrho\, \mathrm{d}h \,, \end{aligned} \tag{6.40}$$

where m_{H} is the mass of a hydrogen atom and ϱ is the density of matter per cm^3 due to hydrogen atoms. By virtue of (6.40), the *coefficient of scattering by free electrons*, referred to unit mass, is

$$s = x\, s_e/m_{\mathrm{H}}\,, \tag{6.41}$$

where s_e is given by the expression (5.70).

Next, in the atmospheres of the stars considered, the true absorption in the observable region of the spectrum is determined largely by *neutral hydrogen*. For this we can use for $\varkappa_\nu$ the expressions (6.1) to (6.3). From these expressions and (6.41) we can write for α_ν (the ratio of the absorption coefficient to the scattering coefficient)

$$\alpha_\nu = \frac{\varkappa_\nu}{s} = \frac{C_0\, p_e}{s_e} \frac{\Psi(\nu, T)}{T^{\frac{3}{2}}}\,. \tag{6.42}$$

And since, from (6.3), the second factor in (6.42) decreases with temperature, it follows that, for some temperature, α_ν will be less than unity for a given frequency and a given p_e, i. e. the part played by electron scattering in the attenuation of the intensity will become even more important than that of the true absorption [though of course $\varkappa_\nu$ will figure as before in Kirchhoff's Law (3.10)]. Moreover, it follows from (6.42) that, for stars with the same effective temperature, the part played by electron scattering is the greater, the less the mean p_e in their photospheres. In other words, electron scattering plays a considerably greater part in the photospheres of supergiants than in those of ordinary stars.

Direct calculations show that the electron scattering may be important even in the photospheres of A-type supergiants. This is confirmed by observations, which show that the discontinuity at the limit of the Balmer series, or rather the quantity D which characterises this discontinuity according to formula (6.28), is smaller in the spectra of early-type supergiants than in those of ordinary stars of the same spectral classes. This is easily explained by means of electron scattering. In fact, any additional source of absorption diminishes the part played by hydrogen, and consequently decreases D.

Electron scattering is particularly important in the photospheres of stars of class O and of hot stars such as, for example, the Wolf-Rayet stars. However, there is unfortunately as yet no sufficiently reliable theory which takes account of the influence of electron scattering on the temperature distribution in the photospheres of stars of classes A, B and O. For Wolf-Rayet stars, which we shall discuss in Part VI, such calculations were first made by V. A. AMBARTSUMYAN [7].

In concluding the consideration, in this part of the book, of stars of early spectral classes, it must be mentioned that, for very hot stars, namely those of class O and the early subdivisions of class B, the absorption of radiation in the ultra-violet region of the spectrum is determined not only by hydrogen, but by *neutral and ionised helium*; see, for example, J.- C. PECKER [121] and A. B. UNDERHILL [167].

To conclude this chapter we may say that, in the cases where the theory of radiative equilibrium has been most developed (the Sun, and stars of classes A 0 to B 2), its conclusions are in satisfactory agreement with the results of observation, as regards the energy distribution in the continuous spectrum and, for the Sun, as regards the law of darkening to the limb. (The introduction of negative hydrogen ions does not noticeably change the law of darkening to the limb obtained for the case of grey material.) Hence it may be assumed that the original postulates of the theory of radiative equilibrium which we have developed are not far from reality, though of course there are a number of problems and discrepancies which require further study. Also, the effect of absorption lines (in whose frequencies part of the photospheric radiation is scattered back) on the radiative equilibrium must be studied. This topic is extremely complicated and has not yet been completely worked out, so that we shall not discuss it here.

Until now we have considered the theory of radiative equilibrium as applied to photospheres with plane-parallel layers. For photospheres whose linear thickness is comparable with the radius of the star (or even greater than it), the application of the theory of radiative equilibrium results in great difficulties. For the case of grey material, the basic physical principles of the theory of the radiative equilibrium of extended photospheres have been worked out by N. A. KOZYREV. In the emergent (observed) radiation of stars with extended photospheres, the relatively cool layers at a relatively small optical depth play an important part (the *extension effect*). Hence the colour temperatures of stars with extended photospheres should be low.

However, in actual stellar atmospheres, the absorption coefficient depends on the frequency. Thus it is necessary to take account of the part played by hydrogen absorption beyond the limit of the Lyman series. Also, for Wolf-Rayet stars the large part played by electron scattering of radiation must be taken into consideration. Moreover, the application of Kirchhoff's Law (3.10) to photospheres of great extent appears insufficiently justified.

Some of the questions of the theory of extended photospheres will be considered in Part VI, which is devoted to stars whose spectra possess bright lines.

Chapter 7. The structure of stellar photospheres

1. The structure of the photospheres of stars of the classes A 0 to B 2. As has been said in the introduction to Part I, the chief problem in the theory of stellar photospheres is to establish the law of variation with depth of the different physical parameters in stellar (or solar) photospheres. In Chapters 4 and 6 we have been concerned with the question of the temperature distribution. However, a knowledge not only of the temperature, but also of the pressure, the density, etc., at every point of the photosphere is necessary for many astrophysical problems. This is particularly important in the construction of a quantitative theory of absorption lines in stellar spectra. As an independent variable we can now take the linear depth in the stellar photosphere, reckoned upwards and downwards from any level; for example, from the level where the temperature T is equal to the effective temperature T_e.

As in the theory of radiative equilibrium, there is here as yet no general and universal method. The solution of the problem is carried out separately for each particular case.

As the simplest example let us consider photospheres of stars of the classes A 0 to B 2. In the photospheres of these stars, the absorption is practically determined by only one element, namely hydrogen. The corresponding theory was first discussed by É. R. Mustel', and later by V. S. Berdichevskaya [20].

The main assumption from which we start is that at every level of the static photospheres which we shall study there exists not only radiative, but also mechanical (hydrostatic) equilibrium. [We may remark that this assumption is valid only on the average, since in stellar photospheres various kinds of local movement of matter must exist. This is indicated by the study not only of the Sun (e.g. granulation), but also of the stars.] The mechanical equilibrium of the stellar photosphere is brought about as a result of the simultaneous action of the following oppositely directed forces: (1) gravity (towards the centre of the star), (2) the gas pressure and radiation pressure (away from the centre of the star).

In order to deduce the equation of equilibrium, we consider a cylinder with base area 1 cm^2 and height $\mathrm{d}h$ considerably less than the linear dimensions of the base; let the axis of the cylinder be normal to the planes of the photospheric layers. The weight of such a cylinder is $g\,\varrho\,\mathrm{d}h$, where g is the acceleration due to gravity at the given point, and ϱ is the total density of matter at the level in question (determined by all atoms). Also, let p be the total gas pressure at this level, and p' the radiation pressure at this level. The equation of equilibrium for the cylinder can be written in the form

$$\mathrm{d}p + \mathrm{d}p' = g\,\varrho\,\mathrm{d}h\,, \tag{7.1}$$

since on penetrating into the star a short distance dh the increase of the gas and radiation pressure by $d(p+p')$ is balanced by a corresponding increase in the weight of the column (with base 1 cm^2) of overlying matter by $g\,\varrho\,dh$. Here it must be remembered that the value of h increases towards the centre of the star.

Let us introduce an expression for dp'. Classical electromagnetic theory tells us that, in the absorption by matter of radiant energy E moving in a given direction, the matter acquires momentum E/c in that direction, where c is the velocity of light. This was first proved experimentally in 1901 by the distinguished Russian physicist P. N. LEBEDEV. In particular, if the atom absorbs a quantum $h\nu$ from a beam of rays, it thereby acquires momentum $h\nu/c$ in the direction of the beam. From this we calculate the radiative impulse acting on the cylinder considered, with base 1 cm^2 and height dh, in the direction of the *normal* to the planes of the layers, since the direction of the force of gravity also coincides with the normal (in the direction to the centre of the star). In particular, there will be no "oblique stresses" due to radiation pressure in the case we are considering, since, as we have already said, the intensity of radiation at any point of the stellar photosphere depends only on the angle θ, and is independent of the azimuth.

The amount of energy absorbed from the beam $I_\nu(\theta)\,d\omega$ by the cylinder in time dt and in the frequency interval from ν to $\nu + d\nu$ is, in accordance with the fundamental definitions of Chapter 2,

$$I_\nu(\theta)\cos\theta\,d\omega\,dt\,.\,\varkappa_\nu\,\varrho\,dh\sec\theta\,, \tag{7.2}$$

where $\cos\theta$ is included to take account of the projection of the unit area [formula (2.3)], and $\sec\theta$ to take account of the oblique course of the beam in the cylinder considered [formula (2.27)]. The momentum given to the cylinder in the direction making an angle θ with the normal is (we disregard the sides of the cylinder in view of the smallness of dh)

$$\frac{I_\nu(\theta)\,d\omega}{c}\,\varkappa_\nu\,\varrho\,d\nu\,dh\,dt\,. \tag{7.3}$$

The component of momentum in the direction of the normal is equal to the expression (7.3) multiplied by $\cos\theta$.

It is also necessary to take account of radiation from the whole sphere, i. e. to effect the integration over ω. This gives

$$\frac{\int I_\nu(\theta)\cos\theta\,d\omega}{c}\,\varkappa_\nu\,\varrho\,d\nu\,dh\,dt = \frac{\pi H_\nu}{c}\,\varkappa_\nu\,\varrho\,d\nu\,dh\,dt\,. \tag{7.4}$$

In order to obtain the force acting on the cylinder, (7.4) must be divided by dt. Next, the result must be integrated over the whole spectrum.

And since the force of radiation pressure acting on the cylinder is the difference of the pressures p' at its lower and upper faces,

$$\mathrm{d}p' = \frac{\pi \int\limits_0^\infty H_\nu \varkappa_\nu \,\mathrm{d}\nu}{c} \varrho \,\mathrm{d}h \,. \tag{7.5}$$

The expression (7.5) is very general. In order to take account of negative absorption (i. e. stimulated emission), the absorption coefficient must be taken in accordance with formula (5.61), i. e. the coefficient of stimulated absorption (corresponding to the coefficient B_{ik}) must be multiplied by $(1 - e^{-h\nu/kT})$. This may be seen from the following considerations.

Let radiation of intensity $I_\nu(\theta)$, moving inside an infinitely small solid angle $\mathrm{d}\omega$, be incident on atoms. Then the effect on these atoms of the beam of rays considered amounts to the following. In absorbing radiant energy the atoms will acquire momenta $h\nu/c$ in the direction of the incident beam. After each absorption, a spontaneous emission of the absorbed quantum will take place (after a certain time), as a result of which the atom will acquire a recoil momentum $h\nu/c$ in the direction opposite to that of the quantum emitted. However, since the re-emission can take place, in general, in any direction, on the average all such recoil momenta will annul one another, and the atom will experience radiation pressure only in the direction of the incident beam.

The process of stimulated emission amounts to this: a quantum $h\nu$ moving inside the beam falls on the atom and stimulates the latter to emit *a similar quantum* $h\nu$ in *the same direction*. The atom thereby emits the quantum in the direction of the beam and acquires a recoil momentum which is always *in the opposite direction*. It is clear that this process diminishes the radiation pressure, in the direction of the beam, on the atoms.

The importance of stimulated emission is quantitatively determined by the ratio of the number of processes of stimulated emission to the number of processes of ordinary (stimulated) absorption, and this ratio should not depend on the angle θ, since in both cases the number of transitions is proportional to the density of radiation. For this reason the resultant momentum communicated to the matter by radiation pressure is proportional to the flux πH_ν.

Let us calculate the above-mentioned ratio of the number of processes. According to (5.47) and (5.50) [where instead of ϱ_ν the expression (5.42) must be substituted], this ratio must be

$$\frac{n_{r+1}\,\beta_\nu v \,\frac{c^3}{8\pi h \nu^3}\, I_\nu(\theta)\, \frac{\mathrm{d}\omega}{c}\, \mathrm{d}n_e}{n_{r,k}\, k_\nu' \,\frac{c\,\mathrm{d}\nu}{h\nu}\, I_\nu(\theta)\, \frac{\mathrm{d}\omega}{c}} \tag{7.6}$$

Using the expressions (5.54), (5.52), (5.11), (5.12), (5.55), (5.49), (5.6) and (5.8), we find that the ratio just given is equal to $e^{-h\nu/kT}$, i. e. the decrease in $\mathrm{d}p'$ when the stimulated emission is taken into account is determined by the factor $1 - e^{-h\nu/kT}$. This is what we wished to prove.

The use of the expression (7.5) to estimate $\mathrm{d}p'$ in the case of photospheres of early-type stars is very laborious, since it would here be necessary to calculate the flux πH_ν for different depths. However, as we shall see, the part played by radiation pressure in this case is very small. Hence we shall not commit a great error if we calculate $\mathrm{d}p'$ for the case of grey material and apply the result obtained to the case of a photosphere with an absorption coefficient depending on the frequency.

For $\varkappa_\nu = \varkappa$, the expression (7.5) takes the form

$$\mathrm{d}p' = \frac{\pi \int\limits_0^\infty H_\nu \,\mathrm{d}\nu}{c} \varkappa \varrho \,\mathrm{d}h = \frac{\pi H}{c} \mathrm{d}\tau = \frac{\sigma T_e^4}{c} \mathrm{d}\tau \; ; \tag{7.7}$$

on the other hand, from (4.23) we obtain by differentiation

$$\mathrm{d}\tau = \tfrac{4}{3} \frac{\mathrm{d}T^4}{T_e^4} \; . \tag{7.8}$$

Introducing (7.8) into (7.7), we find

$$\mathrm{d}p' = \tfrac{1}{3} \frac{4\sigma}{c} \mathrm{d}T^4 = \tfrac{1}{3} a \,\mathrm{d}T^4 \, , \tag{7.9}$$

where $a = 4\sigma/c$ is the density constant of integrated radiation, equal to $7{\cdot}569 \times 10^{-15}$ erg cm^{-3} degree^{-4}. Integrating (7.9), we obtain

$$p' = \tfrac{1}{3} a T^4 \, , \tag{7.10}$$

since $p' = 0$ for $T = 0$, and hence the constant of integration is zero. Thus, for grey material, the radiation pressure in the approximation considered is the same as in the case of strict thermodynamic equilibrium.

Let us return to the equation (7.1). Since, in our case, the absorption of radiation is determined by hydrogen, we can use all the results which we obtained in the last chapter. In particular, we can again introduce the variable ζ defined by the expression (6.5). For this purpose, the equation (7.1) can be rewritten

$$\left(\frac{\mathrm{d}p}{\mathrm{d}T} + \frac{\mathrm{d}p'}{\mathrm{d}T}\right) \frac{\mathrm{d}T}{\mathrm{d}\zeta} \frac{\mathrm{d}\zeta}{\mathrm{d}h} = g \varrho \, . \tag{7.11}$$

For the temperature distribution in the stellar photospheres considered, we can with sufficient accuracy use the equation (6.22), i. e. the equation of the first approximation. For the surface temperature we must take

the value T_0 which is found by means of the method there explained (obtaining of the prescribed flux $\pi H = \sigma T_e^4$ at the boundary of the photosphere). Using (6.22) and (6.5), we have instead of (7.11)

$$\left(\frac{\mathrm{d}p}{\mathrm{d}T} + \frac{\mathrm{d}p'}{\mathrm{d}T}\right) \frac{3\,\sigma\, T_e^4}{4\,\pi\, Z(T)}\, \Phi(p_e, T) = g\, \alpha\,, \tag{7.12}$$

where $\alpha = \varrho/\varrho_{\mathrm{H}}$, and $\varrho_{\mathrm{H}} = m_{\mathrm{H}}\, n_{\mathrm{H}}$, where n_{H} is the number of hydrogen atoms in 1 cm^3. This quantity ϱ_{H} is meant in (6.5). According to Table 1 we can take $\alpha \approx 1.8$.

Our next problem is to transform equation (7.12) into a differential equation for p_e and T, where we shall take p_e as the independent variable. Consequently, we must express the quantities appearing in (7.12) in terms of T and p_e.

Let us first express the gas pressure p in terms of T and p_e. In doing so we take into account the fact that the total gas pressure is practically determined by hydrogen, owing to the extremely high percentage content of the latter. Let 1 cm^3 contain n_{H} hydrogen atoms (both neutral and ionised). Then the partial pressure determined by hydrogen atoms is

$$p_{\mathrm{H}} = n_{\mathrm{H}}\, kT\,. \tag{7.13}$$

In order to obtain the total gas pressure, we must add to p_{H} the partial electron pressure p_e:

$$p_e = n_e\, kT\,, \tag{7.14}$$

where n_e is the number of free electrons in 1 cm^3.

Next, for the temperatures and electron pressures in which we are interested (in the photospheres of A 0 to B 2 stars), the main source of free electrons is hydrogen, as may be calculated from Saha's formula and the p_e found from observation. The part played by electrons liberated by the ionisation of metals is here very small, because of the negligible percentage content of the latter. From these results we can write

$$n_e = x\, n_{\mathrm{H}}\,, \tag{7.15}$$

where x is the degree of ionisation of hydrogen.

From (7.13), (7.14) and (7.15) we find

$$p = p_{\mathrm{H}} + p_e = n_{\mathrm{H}}(1 + x)\, kT\,, \tag{7.16}$$

and also

$$p/p_e = (1 + x)/x\,. \tag{7.17}$$

Further, from (5.16) and (5.17), we can write for hydrogen

$$[x/(1 - x)]\, p_e = K_0\,, \tag{7.18}$$

where

$$K_0 = \frac{u_1}{u_0} \frac{2(2\pi m_e)^{\frac{3}{2}} (kT)^{\frac{5}{2}}}{h^3} e^{-\chi_0/kT} . \tag{7.19}$$

Since the first excitation potential of hydrogen is very large (10·16 eV), the partition function u_0, calculated from formula (5.13), reduces to the first term, even for the temperatures which we are considering. For neutral hydrogen $u_0 = g_{0,1} = 2$, and for ionised hydrogen (protons) $u_1 = 1$.

Consequently we can rewrite (7.19) as

$$K_0 = C\, T^{\frac{5}{2}} e^{-\chi_0/kT} , \tag{7.20}$$

where

$$C = \frac{(2\pi m_e)^{\frac{3}{2}} k^{\frac{5}{2}}}{h^3} . \tag{7.21}$$

Also, it follows from equations (7.17) and (7.18) that

$$p = (p_e^2 + 2 K_0 p_e)/K_0 . \tag{7.22}$$

Differentiating (7.22) with respect to T and taking account of (7.20), we obtain

$$\frac{\mathrm{d}p}{\mathrm{d}T} = \frac{1}{K_0} \left[2 (p_e + K_0) \frac{\mathrm{d}p_e}{\mathrm{d}T} - \frac{p_e^2}{T} \left(\frac{5}{2} + \frac{\chi_0}{kT} \right) \right] . \tag{7.23}$$

Also, since from (7.18)

$$x = K_0/(p_e + K_0) , \tag{7.24}$$

we obtain, according to (6.2),

$$\Phi(p_e, T) = \frac{C_0}{m_{\mathrm{H}}} \frac{K_0 p_e}{T^{\frac{3}{2}} (p_e + K_0)} . \tag{7.25}$$

Using now (7.23), (7.9) and (7.25), we obtain from (7.12)

$$\frac{\mathrm{d}T}{\mathrm{d}p_e} = \frac{(p_e + K_0)\, p_e}{\alpha g \frac{2\pi m_{\mathrm{H}}}{3 C_* \sigma T_e^4} T^{\frac{3}{2}} Z(T) (p_e + K_0) + \frac{p_e^2}{2} \frac{1}{T} \left(\frac{5}{2} + \frac{\chi_0}{kT} \right) - \frac{2}{3} a T^3 p_e K_0} . \tag{7.26}$$

The initial conditions for the integration of this equation are determined by the fact that, at the boundary of the photosphere, where $T = T_0$, the pressure $p_e \equiv 0$ and $\mathrm{d}T/\mathrm{d}p_e = 0$.

The right-hand side of equation (7.26) is a known function of the variables p_e and T. The values of $Z(T)$ are taken from formula (6.23).

The integration of equation (7.26) must be carried out by numerical methods. As a result we find the dependence of T on p_e for given T_e and g. The distribution of the remaining quantities inside the photosphere can be found from the following considerations. From (7.24) and (7.20) we can find the degree of ionisation x for every point of the photosphere. This in turn, by formula (7.17), gives the pressure p at every point of the photosphere. The radiation pressure p' can be calculated from formula (7.10).

To determine the density ϱ we multiply and divide the right-hand side of the equation (7.16) by $\alpha\, m_{\mathrm{H}}$. Then, taking into account the facts that $\varrho = \alpha\, n_{\mathrm{H}}\, m_{\mathrm{H}}$ and that, according to the kinetic theory of gases, $k/m = R/\mu$, we obtain the equation of state

$$p = \frac{R}{\alpha\, \mu_{\mathrm{H}}/(1+x)}\, \varrho\, T\,, \tag{7.27}$$

from which it follows that when the hydrogen becomes ionised the molecular weight decreases, and this leads to an increase in the total gas pressure. Knowing p, T and x, we can calculate ϱ from formula (7.27) for every point of the photosphere.

We now relate all the quantities found to the linear depth in the photosphere. To do this, we eliminate the differential $\mathrm{d}\zeta$ from the equations (6.5) and (6.22), and instead of the function $\Phi(p_e, T)$ we introduce the expression for it given by (6.2). We then obtain

$$\mathrm{d}h = \frac{4\,\pi\, m_{\mathrm{H}}\, \alpha\, Z(T)\, T^{\frac{3}{2}}}{3\,\sigma\, T_e^4\, C_0\, x\, p_e\, \varrho}\, \frac{\mathrm{d}T}{\mathrm{d}p_e}\, \mathrm{d}p_e\,. \tag{7.28}$$

The expression in front of $\mathrm{d}p_e$ on the right-hand side of equation (7.28) is a known function of the pressure p_e, since all the quantities appearing in this equation have already been determined [the derivative $\mathrm{d}T/\mathrm{d}p_e$ is determined in solving equation (7.26)].

The relation between h and p_e and, consequently, the remaining physical parameters, can be established by numerical integration of equation (7.28). As an origin for the measurement of h we can take the level where the temperature T is equal to the effective temperature. We shall reckon the values of h positive into the star from this level, and negative outwards.

2. Discussion of the theoretical results. In Table 4 we give the results of the corresponding calculations, carried out by V. S. Berdichevskaya for $T_e = 10{,}500°$ and $g = 0{\cdot}55 \times 10^4$ cm/sec^2. The last column of Table 4 gives the optical depths, calculated from formula (6.7) with $\lambda = 4700$ Å, for the corresponding levels. It should be noticed that the equation (7.26), from which Table 4 was constructed, contains α

only in the quantity αg. Hence any refinement of the numerical value of α will merely alter the value of g corresponding to this particular solution of equation (7.26).

From Table 4 we can form an idea of the conditions existing in the photospheres of stars of class A 0. We see that the pressures p in these photospheres are very small, of the order of some tens or thousands of bars, while the gas pressure at sea level in the Earth's atmosphere is of the order of millions of bars (one normal atmosphere = 1,013,246 bars). The same is true to an even greater degree when we consider the densities in the photospheres of ordinary stars of class A 0; they are of the order of 10^{-9} to 10^{-11} g/cm³, while the density of air at 0° and 76 cm of mercury is 0·0013 g/cm³.

Table 4

p_e (bars)	$T°$	x	p (bars)	p' (bars)	ϱ (g/cm³)	h (km)	τ_λ[4700 Å]
0	8430	1·000	0	12·86	0		0
5	8432	0·783	11·40	12·88	$1{\cdot}67 \times 10^{-11}$	—4811	—
15	8448	0·556	42·00	12·97	$7{\cdot}03 \times 10^{-11}$	—3210	—
25	8477	0·447	81·00	13·14	$1{\cdot}45 \times 10^{-10}$	—2516	0·005
50	8594	0·350	193·0	13·90	$3{\cdot}66 \times 10^{-10}$	—1662	0·02
100	8890	0·349	386·0	15·91	$7{\cdot}07 \times 10^{-10}$	— 991	0·07
150	9185	0·406	519·0	18·13	$8{\cdot}83 \times 10^{-10}$	— 686	0·13
200	9467	0·480	616·0	20·46	$9{\cdot}67 \times 10^{-10}$	— 491	0·20
300	10,024	0·640	768·0	25·72	$1{\cdot}03 \times 10^{-9}$	— 210	0·34
400	10,607	0·784	912·0	32·25	$1{\cdot}06 \times 10^{-9}$	+ 47	0·54
600	11,746	0·929	1248·0	48·50	$1{\cdot}21 \times 10^{-9}$	+ 611	1·12
780	12,555	0·966	1587·3	63·29	$1{\cdot}41 \times 10^{-9}$	+1102	1·79
1020	13,360	0·982	2059·1	81·15	$1{\cdot}71 \times 10^{-9}$	+1667	2·78

Such small densities and pressures are characteristic not only of the class of stars considered, but in general of the photospheres of the majority of stars. In many cases (photospheres of giants and supergiants) the densities are even several orders of magnitude lower than the figures just quoted, though of course there are also denser photospheres, such as, for instance, those of the white dwarfs. However, the photospheres of the overwhelming majority of stars are very rarefied gaseous configurations.

A glance at the third column of the table shows that, as we penetrate into the photosphere, the degree of ionisation at first decreases, and then begins to increase again. This is easily explained as follows. As is seen from the second column of Table 4, the increase of the temperature in the outermost layers of the photosphere is relatively slow. Hence the change of x in these layers is due to the change of p_e (which increases), and consequently x must decrease in accordance with (7.24). However, on penetrating further into the photosphere, the part played

by the increase of temperature becomes greater and greater; K_0 begins to increase more rapidly than p_e, and the degree of ionisation begins to increase again.

A comparison of the fourth and fifth columns of Table 4 shows that p' is greater than the gas pressure only in the outermost layers of the photosphere, while in the deeper layers it is negligibly small. However, in the outermost layers of the photosphere its influence on the distribution of the various physical parameters turns out to be completely insignificant also. In fact, it follows from equation (7.1) that the equilibrium of the photosphere is determined not by the pressures p and p' themselves, but by their gradients, while it follows from Table 4 that even in the outermost layers of the photosphere $|\,\mathbf{grad}\, p\,| \gg |\,\mathbf{grad}\, p'\,|$. Similar results are obtained for the photospheres of the majority of stars. Consequently, in all such cases the radiation pressure can be neglected. The part played by the latter can become important only in very extended photospheres. Also, selective radiation pressure, i. e. the pressure of radiation in the frequencies of spectral lines, may become important in some cases (very hot stars).

Let us consider the seventh column of Table 4. We are concerned with the layers which mainly produce the continuous spectrum of the star. These are the regions whose optical depth lies, roughly speaking, between 0·1 and 1·5. From regions with $\tau_\lambda > 1{\cdot}5$ the radiation emerges considerably attenuated (owing to the factor $e^{-\tau_\lambda \sec\theta}$). The part played by regions with small τ_λ in producing the emergent radiation is, according to (3.31), proportional to $\Delta\tau_\lambda$ and is consequently small because $\Delta\tau_\lambda$ is small.

It follows from Table 4 that the linear thickness of the layers in question is of the order of some thousands of kilometres. Now the radius of a star of class A 0 with $g = 0{\cdot}55 \times 10^4$ cm/sec^2 is one or two million kilometres. Consequently, we see that in this case the thickness of the star's photosphere is in fact small compared with its radius.

It must be borne in mind that the model here considered of the photosphere of a star of class A 0 is a purely theoretical model, constructed on the assumption that neutral hydrogen is the *only* source of absorption. For values of g greater than that used in the calculations ($5{\cdot}5 \times 10^3$ cm/sec^2), the mean value of p_e will be higher, and negative hydrogen ions will begin to play a significant part. Reasons have been advanced [27] for supposing that the mean electron pressures in the photospheres of stars of class A 0 are of the order of a thousand or even several thousand bars. However, there are other investigations which do not corroborate these conclusions (O. A. Mel'nikov [83], L. H. Aller [2, pp. 196-8]).

3. The application of the law of darkening towards the limb of the star's disc. We have discussed the theory of the structure of stellar photospheres, starting from the fundamental equation of equilibrium (7.1), the temperature distribution found earlier, and other relations. For the Sun, some information on the structure of the photosphere can be obtained from the law of darkening to the limb of the disc, which is found directly from observation. For the expression (3.38) can be regarded as an integral equation determining the law of variation of B_ν with τ_ν, if the law of variation of $I_\nu(\theta, 0)$ with the angle θ is known for a given frequency. Having found this law from observation, we can thus determine also the law of variation of B_ν with τ_ν, which is extremely important for the theory of absorption lines*.

In practice this may be done as follows, following D. Chalonge and V. Kourganoff. We write the equation (3.38) in the form

$$\Phi_\lambda(\theta) = \frac{I_\lambda(\theta, 0)}{I_\lambda(0, 0)} = \int_0^\infty \frac{B_\lambda(T)}{I_\lambda(0, 0)} e^{-\tau_\lambda \sec\theta} \sec\theta \, d\tau_\lambda . \qquad (7.29)$$

The observations show that the function $\Phi_\lambda(\theta)$ can be represented with sufficient accuracy in the following form:

$$\Phi_\lambda(\theta) = A_\lambda + C_\lambda \cos\theta + D_\lambda \cos^2\theta . \qquad (7.30)$$

Introducing this into (7.29), it is easy to prove that the solution of (7.29) is

$$B_\lambda(T)/I_\lambda(0,0) = A_\lambda + C_\lambda \tau_\lambda + \tfrac{1}{2} D_\lambda \tau_\lambda^2 . \qquad (7.31)$$

Thus, having found the coefficients A_λ, C_λ and D_λ from observation, we can find the relation between B_λ and τ_λ in accordance with (7.31). Further, by means of (7.31) we can study the variation of τ_λ with λ, and also of $\varkappa_\lambda$ with λ, etc., for every level in the solar photosphere (i. e. for every value of T). This method of studying the solar photosphere may be called the "optical" method. The value of this method was first remarked by P. Parkhomenko in 1934. The method in question was also first used by her; in recent years it has come into very general use. The coefficients A_λ, C_λ and D_λ have been obtained by R. Canavaggia, D. Barbier and others from the most recent observational material.

It must be remarked that the accuracy with which the dependence of B_λ on τ_λ is found by this method depends very strongly on the accuracy with which the function Φ_λ is determined at the limb of the Sun's disc.

* The same can be done, with less certainty, for some eclipsing variables also.

Chapter 8. The application of the laws of thermodynamic equilibrium to stellar photospheres

1. **Preliminary remarks on the deviations from thermodynamic equilibrium in stellar photospheres.** In the preceding chapters we have several times made use of laws which are, strictly speaking, correct only for the case of thermodynamic equilibrium. Let us enumerate these laws: (1) the Maxwellian distribution of the velocities of atoms and electrons [for electrons cf. formula (5.52)]; (2) Boltzmann's formula (5.9), which gives the distribution of atoms among the quantum levels; (3) Saha's ionisation formula (5.11); (4) the Kirchhoff-Planck law (3.10). The validity of the application of these laws to stellar photospheres is by no means evident. Hence this question needs special consideration.

However, before going on to a detailed discussion of the laws which we have enumerated, we must ascertain why these laws may in general be violated in stellar photospheres. In other words, we must find out how the physical conditions existing in stellar photospheres differ from those which hold in thermodynamic equilibrium.

The first violation of the conditions which characterise strict thermodynamic equilibrium arises from the existence of a temperature gradient in stellar photospheres. The variation of temperature in the photosphere leads to the dependence of the intensity of radiation on the direction [see, e. g., formulae (3.36) and (3.37)]. This can be explained as follows. Let the intensity of radiation of frequency ν going in a given direction be $I_\nu(\theta,\tau_\nu)$. Then we can numerically equate this intensity to the Planck intensity B_ν and determine from this equation the temperature T, which will of course be a function of depth, of frequency and, what is here particularly important, of direction; we denote it by $T(\theta,\nu)$:

$$I_\nu(\theta,\tau_\nu) = \frac{2h\nu^3}{c^2} \cdot \frac{1}{\exp[h\nu/kT(\theta,\nu)] - 1} \cdot \tag{8.1}$$

Conversely, the substitution of the temperature $T(\theta,\nu)$ in the Planck function gives the numerical value of the intensity of ν-radiation moving in a given direction θ. The temperature $T(\theta,\nu)$ thus introduced may be called the **temperature of the radiation** (or **brightness temperature**). In what follows we shall denote it by T_E. Contrary to what is the case in stellar photospheres, the radiation in conditions of thermal equilibrium is strictly isotropic. In other words, the value of $T(\theta,\nu)$ is the same in all directions.

The effect considered, which we may call the *geometrical effect*, is greatest at the boundary of the photosphere, where for $\frac{1}{2}\pi < \theta \leq \pi$ the intensity $I_\nu(\theta,0) = 0$; thus in this range $T(\theta,\nu) = 0$. For the outermost parts of the photosphere (conventionally called the reversing

layer) and for the envelopes of the star which lie still further out, the deviations from equilibrium radiation which are due to the geometrical effect are allowed for by introducing the **dilution factor** W. We suppose that the intensity of the continuous radiation emerging from the star is independent of direction and is given by Planck's formula with temperature T. Then (Fig. 19) the density of radiation at any point at a distance r from the centre of the star is equal, by (5.37), to

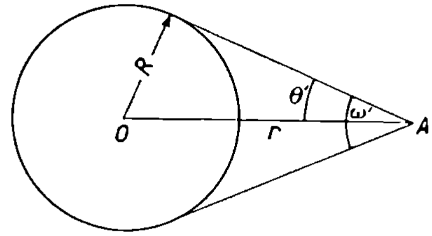

FIG. 19

$$\varrho_\nu = \frac{1}{c}\int_{\omega'} B_\nu(T)\,\mathrm{d}\omega = \frac{4\pi}{c} B_\nu(T)\frac{\omega'}{4\pi}, \tag{8.2}$$

where ω' is the solid angle subtended by the star at the point A. It follows from Fig. 19 that the ratio of this solid angle to 4π, which we shall call the *dilution factor*, is

$$W = \frac{\omega'}{4\pi} = \frac{2\pi\int_0^{\theta'}\sin\theta\,\mathrm{d}\theta}{4\pi} = \tfrac{1}{2}\left[1-\sqrt{\{1-(R/r)^2\}}\right]. \tag{8.3}$$

Introducing the expression for $B_\nu(T)$ into (8.2), we obtain

$$\varrho_\nu = W\frac{8\pi h\nu^3}{c^3}\frac{1}{\exp(h\nu/kT)-1}. \tag{8.4}$$

On comparing (8.4) with (5.53), we see that in the case considered the density of radiation at the given point A is equal to the equilibrium density of radiation at temperature T, multiplied by the dilution factor W. In particular, it follows from (8.3) that in the outer layers of the photosphere, where $R \approx r$, $W \approx \frac{1}{2}$; this is obvious, since here the radiation comes only from below.

Thus the factor W is a *measure of the deviation of the density of radiation from the equilibrium density*. For strict thermodynamic equilibrium at a temperature T, the radiation at the point A comes from all directions, and the dilution factor $W = 1$.

The second reason for the violation of thermodynamic equilibrium in stellar photospheres is the non-Planckian character of the radiation. In fact, even for a given point and a given direction, the variation of the intensity of radiation $I_\nu(\theta,\tau_\nu)$ with frequency may differ greatly from the Planckian in many cases. This may be seen, for example, from the graph constructed to show the radiation emerging from a star (Fig. 17). This figure is constructed for the mean intensity over direction, $\bar{I}_\lambda$. However, it is clear that a similar dependence of I_λ on λ holds good

for the various angles θ also. The non-Planckian variation of the intensity of radiation with frequency is characteristic not only of the actual boundary of the star, but also of the deeper layers of the stellar photosphere. Only as $\tau_\nu \to \infty$ does the value of $I_\nu(\theta)$ tend to the Planckian for all directions [cf. (6.20)].

The deviations from Planck radiation can also be shown by the fact that the temperature $T(\theta, \nu)$ in stellar photospheres, defined by the relation (8.1), depends on the frequency ν, even for a given point and a given direction. Thus, for example, if we return to Fig. 17, the value of $T(\theta, \nu)$ — in this case for the boundary of the photosphere — is anomalously high just before the limit of the Balmer series (segment AB), and anomalously low just beyond this limit (segment CD).

The third factor which brings about the violation of thermodynamic equilibrium in stellar photospheres is the difference of the *kinetic temperature* T_k from the temperature of the radiation. By the kinetic temperature we shall mean the temperature which appears in the formula for the Maxwellian law of distribution of velocities. In an ideal gas this temperature is the same as the temperature which appears in the equation of state:

$$p = nkT\,. \tag{8.5}$$

The difference of T_k from the temperature of the radiation in stellar photospheres is quite evident merely from the circumstance that at any point of the photosphere the temperature of the radiation is a function of the angle θ and the frequency. Consequently, an equality between T_k and $T(\theta, \nu)$ can exist only for some single direction and some particular frequency ν (or at most for a few). In general this equality does not hold at all. It must be remarked here that the discovery has recently been made of the operation in stellar atmospheres of various factors which raise the kinetic temperature quite independently of the radiation field. Thus, for example, it is known that in the solar corona the kinetic temperature of the electrons is of the order of a million degrees, while the intensity of solar radiation is relatively small.

Finally, it must be borne in mind that excitation of the atoms by radiation takes place in the frequencies of spectral lines, these being absorption lines or (in some cases) emission lines. In the first case this decreases the temperature of the radiation in the frequencies concerned; in the second, it increases it.

Summarising what has been explained above, we can say that the radiation field in stellar photospheres is extremely complex, and cannot be described by any one value for the temperature of the radiation. All this implies a violation of the conditions which hold for thermodynamic equilibrium, where the temperature appearing in *all* the thermal

laws and, in particular, in those which we have enumerated, has one and the same value. Thus the question of the application of the laws of thermal equilibrium to stellar photospheres is extremely important.

In the present chapter we shall consider only three of the laws which we enumerated at the beginning. We shall be able to deal with the application of Boltzmann's Law only after having discussed the theory of absorption lines.

2. **The velocity distribution of atoms and electrons in stellar photospheres.** Let us first consider the question of the Maxwellian distribution of velocities. All the data indicate that the Maxwellian distribution of velocities is maintained with entirely sufficient exactness, not only in relatively dense photospheres, but also in such rarefied formations as the planetary nebulae and the interstellar gas.

In general there are two fundamental factors which might cause deviations from the Maxwellian distribution: (1) the presence of electric fields; (2) the difference of the ionising radiation from the equilibrium radiation for which the law (5.53) holds.

Let us deal with the first factor. The presence of weak electric fields on the surface of the Sun has been established beyond doubt. In particular, this is indicated by the study of the motions of prominences. There must be such fields also in the atmospheres of other stars. The study of the influence of electric fields on the distribution of velocities in a gas forms the subject of a special branch of physics — the physics of gas discharges. One of the principal divisions of gas discharge physics is the subject of *plasma*. **Plasma** is the name given to a highly ionised gas which is characterised by the fact that no noticeable volume charges are formed in it. This gas has a high conductivity (because of the presence of a large number of free electrons). The gas in stellar atmospheres (and, for instance, in the Earth's ionosphere) is such a plasma. It is true that the study of prominences shows that volume charges are apparently present in this particular case. However, these charges are undoubtedly small. The positive charge of the entire solar envelope, brought about by the thermal dissipation of the easily "volatilised" electrons, is even smaller. If a considerable charge is formed anywhere, it is rapidly dispersed on account of the very high conductivity of the photospheric gas. Hence the photosphere may be considered as a neutral configuration.

Numerous laboratory investigations have shown that, as a rule, the electrons of a plasma have a Maxwellian distribution of velocities; this indicates an intense interaction (collisions) between the electrons of the plasma. Deviations from a Maxwellian distribution of electrons in a plasma may occur when their concentration is small and, in conse-

quence, collisions between them are rather rare. However, in stellar atmospheres, where the ionisation of the gas is fairly high because of its high temperature, the exchange of energy between electrons will also be fairly considerable. Thus there is every reason to suppose that the effect of the electric fields in stellar photospheres on the distribution of electron velocities can be neglected.

Let us now turn to the second factor. The absence of a Maxwellian distribution of velocities of the electrons might be expected on account of the fact that free electrons arise mainly from the photo-ionisation of atoms (see below), and the distribution of the ionising radiation may deviate very markedly from the equilibrium distribution (5.53). Nevertheless, this factor too cannot be at all important. If these electrons were captured by ions immediately after the photo-ionisation (recombination processes), then of course the velocity distribution of the electrons at any moment would reflect the distribution in frequency of the ionising radiation and the deviation of this radiation from isotropy in direction. Consequently, in some cases (cf., for example, Fig. 17) one might expect a very marked deviation from the Maxwellian distribution. This, however, is by no means the case. Elementary calculations show that, under the conditions in stellar atmospheres, where there is a sufficient number of free electrons by virtue of the relatively high ionisation of the matter, the probability of the *recombination* of an electron with an ion is several orders of magnitude less than the probability of the *collision* of two free electrons. This is due to two circumstances. Firstly, the effective cross-section for recombinations [the quantity β_ν in formula (5.48)] is very small, and in every case some orders of magnitude smaller than the effective cross-section for the collision of two electrons. Secondly, collisions between electrons are, by virtue of their relatively small mass and consequent relatively large mobility, considerably more frequent than collisions between electrons and ions.

It follows from what has been said that an electron which has been removed from an atom by photo-ionisation immediately undergoes a very large number of collisions with other free electrons. Consequently, the distribution of velocities which corresponds to a non-equilibrium density of the ionising radiation will be completely "erased" as a result of collisions, and replaced by a Maxwellian distribution of velocities. Here it is clear that the considerations given are unconnected with the density of matter, and are applicable also to such rarefied objects as planetary nebulae and the interstellar gas, where the ionisation of atoms again produces a fairly large number of free electrons. It follows from this that the second factor which might destroy the Maxwellian distribution of velocities of electrons (the non-equilibrium density of the ionising radiation) is also unimportant.

The question of the velocity distribution of heavy particles (atoms and ions) is more complex. However, similar conclusions can also be arrived at here. In fact, in consequence of the quite considerable exchange of energy between these particles and the electrons, a Maxwellian distribution of velocities must hold in this case also. In this connection, the loss of the energy communicated to the heavy particles by the electrons will be very small in view of the low thermal conductivity of the stellar photosphere, and so will the energy which is lost by the heavy particles in excitation processes, since the efficiency of such processes is generally small. All this means that we are quite justified in expecting equality between the kinetic temperature of the electrons and that of the heavy particles (i. e. the temperature which appears in the corresponding formula for the Maxwellian distribution).

3. **The photo-ionisation of atoms.** We pass now to a consideration of Saha's formula. However, before investigating its applicability to stellar photospheres, we shall make a number of general remarks relating to ionisation and recombination processes, since the interconnection of these processes determines the state of ionisation.

The ionisation of atoms may take place in two ways: A_{ion} — ionisation of atoms *by radiation* (photo-ionisation), and B_{ion} — ionisation as a result of *collisions* of atoms with electrons or other particles, for instance neutral atoms or ions.

It is true that, besides this, there exists a third possibility — the ionisation of atoms by *interatomic electric fields* which are produced as a result of the passage of charged particles close to atoms. However, this ionisation plays an important part only for highly excited atoms, of which there are relatively few, by reason of the Boltzmann distribution. Hence the efficiency of the third ionising factor is negligibly small under the conditions in stellar photospheres, and for this reason we shall not discuss it here.

The processes converse to the two above-mentioned ionisation processes are respectively the recombination processes: A_{rec} — recombination of two particles *with emission of a quantum $h\nu$*, and B_{rec} — recombination of two particles *without emission of radiant energy*, but with a third particle taking part, to which the excess energy is given. We can regard the latter recombination process as a *triple collision of the second kind*.

Let us compare the efficiency of the two processes A_{ion} and B_{ion}. To do so we first consider more closely the subject of photo-ionisation. We calculate the number of photo-ionisations Z_{pi} which take place in 1 cm³ in 1 second, relating to atoms of some given kind. In doing so we must take into account that the photo-ionisation of atoms can take place from any level, so that, strictly speaking, it is necessary

to take all the levels in the atom into consideration. Let the atomic absorption coefficient relating to the photo-ionisation of the atom from the level k be $(k_\nu')_k$. Then, according to (5.46) and (3.15), the number of photo-ionisations from the level considered is

$$\frac{4\pi}{h} \int_{\nu_k}^{\infty} J_\nu \frac{(k_\nu')_k}{\nu} n_{r,k} \, d\nu , \tag{8.6}$$

where $h\nu_k = \chi_{r,k}$ is the binding energy of the given level, since by (5.8) the ionisation begins only at the frequency which corresponds to the detachment of the photoelectron ($v \equiv 0$).

Since the photo-ionisation takes place from all levels, we can write for the quantity Z_{pi} required

$$Z_{\text{pi}} = \frac{4\pi}{h} \sum_k \int_{\nu_k}^{\infty} J_\nu \frac{(k_\nu')_k}{\nu} n_{r,k} \, d\nu , \tag{8.7}$$

where the summation is over all possible values of k. It follows from (8.7) that the quantity Z_{pi} is determined mainly by the spectral regions where the product $J_\nu (k_\nu')_k \, n_{r,k}$ is large. Let us consider a case where the deviation of J_ν from the equilibrium value is especially large; for instance, the ionisation of hydrogen atoms at the surface of a star of class A 0. For the boundary of the star, where the incident intensity is zero, we can write, from (3.50), (3.51) and (4.47),

$$(J_\nu)_0 = \tfrac{1}{2} (I_\nu)_0 = \tfrac{1}{2} \bar{I}_\nu , \tag{8.8}$$

where the values of $\bar{I}_\nu$ may be taken from Fig. 17, converting $\bar{I}_\lambda$ into $\bar{I}_\nu$ according to the formulae (4.36) and (4.38). The values of $n_{r,k}$ may be taken in accordance with the Boltzmann distribution (5.27) for $T = 8430°$ (see Table 3), putting $n \equiv k$, and the values of $(k_\nu')_{k=n}$ from formula (5.24). The value of $n_{r,1} = n_{0,1}$ is immaterial; we are concerned only with the relative variation of $J_\nu (k_\nu')_k \, n_{r,k}$ with frequency.

In using Fig. 17 and Boltzmann's formula (5.27), we have postulated the validity of the formulae of thermodynamic equilibrium. However, this need not perturb us, since here it is possible to use the method of successive approximations. Moreover, the theory of radiative equilibrium which we have considered is in satisfactory agreement with observation, so that in the first approximation the application of the formulae of thermodynamic equilibrium to stellar photospheres may be considered sufficiently valid. The dependence of $J_\nu (k_\nu')_k \, n_{r,k}$ on fre-

quency for the case in question is shown graphically in Fig. 20. We see that the value of Z_{pi} is mainly determined by the regions of the spectrum where the product $(k'_\nu)_k\, n_{r,\,k}$, that is, the absorption coefficient refered to 1 cm^3 or to unit mass, is large. The effect on Z_{pi} of changes in J_ν is much less important; where $(k'_\nu)_k\, n_{r,\,k}\, J_\nu$ is large, the value of J_ν has actually diminished (this is true for stars of other spectral classes also).

On the other hand, where $n_{r,\,k}(k'_\nu)_k$ is large, the intensity of radiation at a given point is close to the Planck intensity corresponding to the temperature at this point. In fact, we have already remarked more

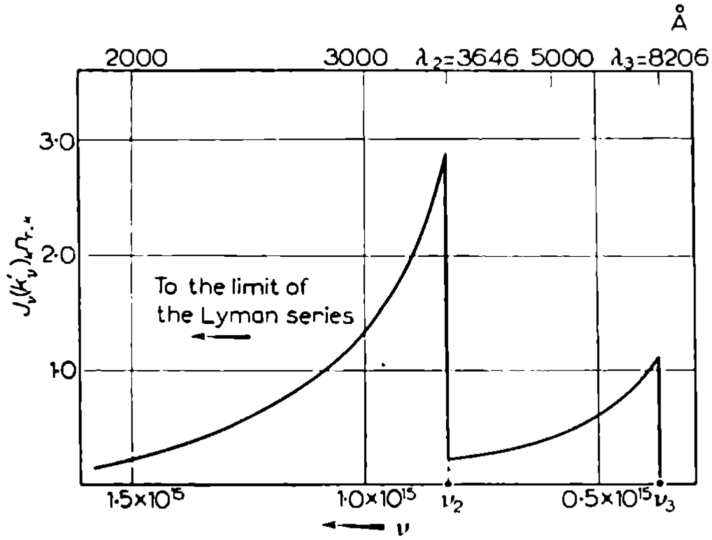

FIG. 20

than once that the intensity of radiation at any point (for $\theta \geq \frac{1}{2}\pi$) is a weighted mean of the values of B_ν along the given direction. (For $\theta < \frac{1}{2}\pi$ this statement is correct for large τ_ν.) And if the opacity of the photospheric matter is large for radiation of the frequency concerned [i. e. $n_{r,\,k}(k'_\nu)_k$ is large], the radiation at the point considered will come from regions lying close to this point, i. e. will have a characteristic temperature which is little different from the temperature at that point. [So far we are assuming the validity of Kirchhoff's Law (3.10), and by the temperature at the point we understand the temperature which appears in (3.10).]

Thus we arrive at an important result. The photo-ionisation of atoms in the case we are considering is mainly determined by the *intensity of radiation of a black body having the same temperature as that at the point in question.*

However, the application of this result to other cases needs a number of reservations. Firstly, the fact that the photo-ionisation is determined mainly by the intensity of radiation given by Planck's Law for the

temperature of the point considered does not mean that the ionising radiation is equilibrium radiation. In fact, it is necessary to take account also of the geometrical effect in the outer layers; we have discussed this above. Both at the extreme boundary of the star (where $I_\nu' \equiv 0$) and outside the star, this effect can be taken into account by means of the dilution factor W. Since, at the boundary of the photosphere, according to our result, the intensity of radiation, in the frequencies which are the most important as regards ionisation, is approximately $B_\nu(T_0)$, where T_0 is the surface temperature (the local temperature at this level), we can approximately write (8.7) as

$$Z_{\rm pi} \approx \frac{4\pi}{h} W \sum_k \int\limits_{\nu_k}^{\infty} B_\nu(T_0) \frac{(k_\nu')_k}{\nu} n_{r,k} \, d\nu \,. \tag{8.9}$$

For elements of which the content in the photosphere is fairly high, the value of J_ν approximates to the equilibrium value $B_\nu(T)$ for the level in question, even at not very large depths in the photosphere. Calculations show that the photo-ionisation of atoms takes place chiefly by the detachment of an electron from the ground level, since an overwhelming proportion of the atoms are in the ground state. For this reason the product $(k_\nu')_k \, n_{r,k}$ is always relatively large for the ground level ($k = 1$). Consequently, the opacity of the material for the "ground continuum" (formed by the ionisation of atoms from the ground level) is relatively large, and for even a small depth we have $\tau_\nu > 1$. Therefore, in accordance with (6.20) and (6.21), J_ν will not differ markedly from B_ν.

An estimate of the value of τ_ν can be made from the following considerations. We expand the Planck function $B_\nu(T)$ in a series of powers of the optical depth and neglect all powers of t_ν above the first. This gives

$$B_\nu(T) = a_\nu + b_\nu t_\nu = B_\nu(T_0) + \left(\frac{dB_\nu}{dt_\nu}\right)_{t_\nu=0} t_\nu \,. \tag{8.10}$$

This expansion may lead to very serious errors in some cases. However, in the regions where the absorption is large, the expression (8.10) is sufficiently accurate for a calculation of the intensity, since here $B_\nu(T)$ changes relatively little over the range of optical depth which determines the intensity. Introducing (8.10) into (3.63) and (3.64), we find for I_ν and I_ν'

$$I_\nu(\tau_\nu) = B_\nu(\tau_\nu) + \tfrac{2}{3} b_\nu \,, \tag{8.11}$$

$$I_\nu' = B_\nu(\tau_\nu) - \tfrac{2}{3} b_\nu + [\tfrac{2}{3} b_\nu - B_\nu(T_0)] \, e^{-\frac{3}{2}\tau_\nu} \,. \tag{8.12}$$

We consequently obtain for J_ν, according to (3.51),

$$J_\nu = \tfrac{1}{2} (I_\nu + I_\nu') = B_\nu(T) + \tfrac{1}{2} [\tfrac{2}{3} b_\nu - B_\nu(T_0)] \, e^{-\frac{3}{2}\tau_\nu} \,. \tag{8.13}$$

Next, for layers of the photosphere which have a large opacity in the frequencies considered, the value of b_ν is always less than $B_\nu(T_0)$. In fact, for the boundary of the star we have from (8.11)

$$I_\nu(0) = B_\nu(T_0) + \tfrac{2}{3}\, b_\nu\,. \tag{8.14}$$

As we have already remarked, when the opacity of the layers is large, $I_\nu(0) \approx B_\nu(T_0)$, and what we have just said confirms this. It then follows from (8.13) that, even for $\tau_\nu \approx 2$, the value of J_ν is very close to that of B_ν. Quantitative calculations show that, for elements of which the content in the photosphere is fairly high, this value of τ_ν (for the ground continuum) is reached in the outermost layers of the photosphere. Consequently, the value of J_ν for the ionising radiation is, almost throughout the photosphere, close to $B_\nu(T)$, where T is the temperature at the point considered.

The second reservation which we must make is as follows. All the results just obtained refer to the most abundant elements in the stellar photosphere (H, He, etc.). For elements whose percentage content (by number of atoms) is relatively small, the optical thickness τ_ν due to them may be very small, even for the ground continuum. Hence each case needs additional investigation. However, it is to be noted that very frequently, apart from the dependence on the content of the given element in the photosphere, the photo-ionisation is determined by the temperature at the point considered. Let us consider, for example, the spectrum of the Sun. In the ultra-violet region of the Sun's spectrum beyond $\lambda = 3000$ Å, there is a concentration and overlapping of a great number of ground continua belonging to the metals, and consequently the intensity of radiation at every point is close to $B_\nu(T)$, where T is the temperature at the given point. Hence, if the ground continuum of any element whose relative content is low falls in this spectral region, then we have also, for this element, $I_\nu \approx B_\nu(T)$.

Finally, it must be borne in mind that in the ultra-violet region of the spectrum there is a concentration of very many resonance absorption lines of various elements, and this causes a deviation of J_ν from B_ν.

Summarising what has been explained above, we may draw the following conclusions: the photo-ionisation of elements with a relatively high content a_s (and in many cases with low a_s) is determined mainly by the intensity $B_\nu(T)$, where T is the temperature at the point for which $Z_{\rm pi}$ is calculated. Almost throughout the photosphere we can take $J_\nu \approx B_\nu(T)$ for these elements, while at the boundary the formula (8.9) can be used, with the dilution factor $W \approx \frac{1}{2}$. In the remaining cases (for instance, in the intermediate surface layers where the transition takes place from the formula $J_\nu = WB_\nu$ to the formula $J_\nu = B_\nu$),

it is necessary to return to the fundamental general formula (8.7), where we can restrict ourselves, for the purpose of estimating Z_{pi} in the first approximation, to photo-ionisation from the ground level alone.

4. **The ionisation of atoms by collision. Comparison with photo-ionisation. Recombination processes.** Let us now estimate the number of ionisations *by collision* (B_{ion} processes). Let us consider first the ionisation of atoms and ions *by electrons*. We can regard the atoms and ions as stationary compared with the electrons, which have larger velocities on account of their smaller mass. Hence we can use formula (5.48) here also. If $q_{r,k}(v)$ is the effective cross-section for the ionisation of an r times ionised atom in the state k by electrons whose velocity is between v and $v + \mathrm{d}v$, then the number of ionisations which take place in 1 cm^3 in 1 second as a result of these collisions is

$$n_{r,k}\, q_{r,k}(v)\, v\, \mathrm{d}n_e\,, \tag{8.15}$$

where the effective ionisation cross-section depends on the velocity v of the passing electron, as is shown by both theory and experiment. We can take for $\mathrm{d}n_e$ the Maxwellian distribution of velocities (5.52), in view of what has been said at the beginning of the present chapter.

Since the ionisation of the atom from the kth level can be effected by any electron whose kinetic energy is greater than or equal to the binding energy $\chi_{r,k}$, the total number of ionisations from the kth level by collisions is

$$n_{r,k}\, n_e\, 4\,\pi \left(\frac{m_e}{2\,\pi\, kT}\right)^{3/2} \int_{v_k}^{\infty} q_{r,k}(v)\, v^3\, e^{-m_e v^2/2kT}\, \mathrm{d}v\,. \tag{8.16}$$

The value of v_k is determined, according to the relation mentioned, by

$$\chi_{r,k} = \tfrac{1}{2}\, m_e\, v_k^2\,. \tag{8.17}$$

And since the ionisation of atoms can take place from *any* level, the total number of ionisations as a result of collisions with electrons is

$$Z_{\text{cl}} = n_e\, 4\,\pi \left(\frac{m_e}{2\,\pi\, kT}\right)^{3/2} \sum_k n_{r,k} \int_{v_k}^{\infty} q_{r,k}(v)\, v^3\, e^{-m_e v^2/2kT}\, \mathrm{d}v\,. \tag{8.18}$$

In order to compare Z_{pi} with Z_{cl}, it is necessary to know the value of $q_{r,k}(v)$. Here it must be noted that the average thermal velocities of the electrons in stellar photospheres are relatively small. The energies corresponding to them are also relatively small. In Table 5 we give, for a number of temperatures, the root-mean-square velocities v_t and the energies of the electrons in electron-volts.

The calculations are carried out in accordance with the well-known formulae

$$\tfrac{1}{2} m_e v_t^2 = \tfrac{3}{2} kT = V_0 e/300 , \tag{8.19}$$

where V_0 is expressed in electron-volts.

Table 5

$T°$	v_t (km/sec)	V_0 (eV)	$T°$	v_t (km/sec)	V_0 (eV)
3000	369	0·39	20,000	953	2·59
5000	477	0·65	30,000	1168	3·88
7000	564	0·90	50,000	1508	6·47
10,000	674	1·29	75,000	1847	9·70
15,000	826	1·94	100,000	2132	12·93

It follows from this table that the mean energy of the electrons amounts to a few electron-volts. It is true that the electrons with velocities $v > v_k$ are those which participate in the ionisation of the majority of atoms. However, there are so few electrons with energies of some tens of electron-volts that they play a negligibly small part in this respect. Thus in ionisation by collisions the slow electrons are those chiefly concerned. Unfortunately, the data on the values of $q_{r,k}(v)$ for this case are very meagre. They do not exist for hydrogen nor for practically all the other principal chemical elements found in stellar photospheres. In Fig. 21 we give the cross-section $q_{0,1}(v)$ for the ionisation of neutral helium from the ground level [called $q(v)$ in the figure]. As the unit of measurement of $q_{0,1}(v)$, the value πa_0^2 is taken, where a_0 is the radius of the first Bohr orbit of the hydrogen atom: $a_0 = 0{\cdot}529 \times 10^{-8}$ cm. The energy of the ionising electron in electron-volts is placed on the axis of abscissae.

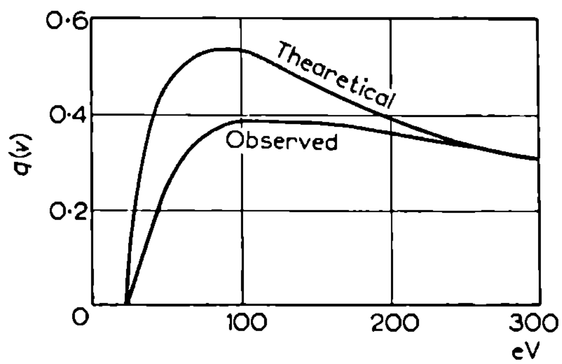

FIG. 21

In the majority of cases, $q_{r,k}(v)$ may be considered independent of the velocity and some *mean* value $\bar{q}_{r,k}$ may be taken for this quantity. On the basis of existing data it can be supposed that the values of $\bar{q}_{r,k}$ for various elements and various states of ionisation and excitation vary between 10^{-17} and 10^{-16} cm². If we introduce the mean value $\bar{q}_{r,k}$ instead of $q_{r,k}$ in equation (8.16), putting

$$m_e v^2/2\,kT = x \tag{8.20}$$

and effecting the elementary integration, we obtain the formula for the total number of ionisations by electron collision from the kth level:

$$n_{r,k}\, n_e\, \bar{q}_{r,k} \sqrt{\left(\frac{8\,kT}{\pi\, m_e}\right)} \left(1 + \frac{\chi_{r,k}}{kT}\right) e^{-\chi_{r,k}/kT} \tag{8.21}$$

where in general it must be remembered that the temperature which appears here is the *kinetic temperature of the free electrons*; this we shall distinguish where necessary by the suffix ε.

Let us now turn to the ionisation of atoms and ions *by collisions with heavy particles*, i. e. by collisions with atoms and ions. The existing experimental data, as well as theory, show that the efficiency of the ionisation of atoms and ions by heavy particles is much less than that of the ionisation by electrons. The ionisation by heavy particles can become important in comparison with the ionisation by electrons only in the case where the number of free electrons in 1 cm^3 is several orders of magnitude smaller than the total number of heavy particles in the same volume. It is possible that such a situation exists in the atmospheres of the cooler stars, where n_e is considerably less than the number of atoms of neutral hydrogen in 1 cm^3. However, we have not yet at our disposal the necessary data for quantitative results, namely, the effective cross-sections for the two components, which are still unknown.

Ionisation by collisions is, in the great majority of cases, of minor importance compared with ionisation by radiation. To show this, let us consider, for instance, the ionisation of hydrogen in the photospheres of stars of class A 0 at the level where $T \approx 10{,}000°$. For the reasons already stated, we can confine ourselves to the consideration of ionisation from the ground level. For the absorption coefficient, neglecting the dependence of g' on frequency, we can use the expression (5.24), rewriting it as

$$(k_\nu')_1 = k_1 (\nu_1/\nu)^3\,, \tag{8.22}$$

where ν_1 is the frequency of the limit of the Lyman series, so that $h\nu_1 = \chi_0$. The value of k_1 is $6{\cdot}3 \times 10^{-18}$ cm^2. Next, since the optical depth just beyond the series limit is considerably greater than unity at the depth where $T \approx 10{,}000°$, we can with good accuracy put $J_\nu = B_\nu$ (10,000°) in (8.6). Thus the ratio of the number of ionisations by electron collision to the number of photo-ionisations (in 1 cm^3) is, by (8.6) and (8.21),

$$\alpha = \frac{n_{0,1}\, n_e\, \bar{q}_{0,1} \sqrt{\left(\frac{8\,kT_\varepsilon}{\pi\, m_e}\right)} \left(1 + \frac{\chi_0}{kT_\varepsilon}\right) e^{-\chi_0/kT_\varepsilon}}{\frac{4\pi}{h} \int\limits_{\nu_1}^{\infty} B_\nu \frac{(k_\nu')_1}{\nu}\, n_{0,1}\, d\nu} \tag{8.23}$$

where we have especially noted that the numerator is a function of the electron temperature at the given level. Since also, for the temperature considered, $e^{h\nu/kT} \gg 1$ for all $\nu \geq \nu_1$, we can neglect unity in the denominator of Planck's function (3.11). Introducing (3.11) and (8.22) into (8.23), and using also (6.35) and (5.12), we obtain

$$\alpha = \frac{\bar{q}_{0,\,1} \sqrt{\left(\frac{8}{\pi m_e}\, k T_\varepsilon\right)} \left(1 + \frac{\chi_0}{k T_\varepsilon}\right) e^{-\chi_0/k T_\varepsilon}}{\frac{8 \pi k_1 \nu_1^3}{c^2}\, E_1\, (\chi_0/kT)}\, p_e \,. \tag{8.24}$$

Assuming that the temperature T is equal to the electron temperature T_ε (see below), we shall take for their value 10,000°, for $\bar{q}_{0,\,1}$ the value 10^{-17} cm^2, and for p_e, from Table 4, the value 300 bars; for these values, $\alpha \approx 5 \times 10^{-3}$. Thus the part played by ionisation by electron collision is indeed vanishingly small in comparison with that of photo-ionisation. The part played by heavy particles — mainly hydrogen atoms — in this case is also negligible. In fact, the values of x given in Table 4 indicate that here n_e is comparable with n_{II}, the total number of hydrogen atoms in 1 cm^3. At the same time, we have already said that the efficiency of ionisation by heavy particles is several orders of magnitude less than that of ionisation by electrons, i. e. for heavy particles the value of α is even less than the figure mentioned.

For the ionisation of hydrogen in the solar photosphere at the level where $T = T_e = 5700°$ and p_e is of the order of 30 bars (see Part III), the value of α given by formula (8.24) is 10^{-4}. It is true that here n_e is approximately three orders of magnitude less than n_{II}. However, the ionisation by heavy particles can be completely neglected even in this case, because of its inefficiency.

Similar calculations lead to analogous results for other elements. In general, the ionisation by collisions is negligibly small in the majority of ordinary stellar photospheres, in comparison with photo-ionisation. The physical reason for this is the relatively low density of the matter which forms stellar photospheres. Hence it is quite possible that in some cases, for example in the dense photospheres of white dwarfs, the value of α is greater than unity. However, this question has not yet been wholly clarified from the quantitative aspect.

In conclusion, it must be remarked that, in the outer envelopes of stars (for example in the higher layers of the solar chromosphere and in the solar corona), the value of T_ε in formula (8.24) may be considerably higher than that of T. As may be easily calculated from formula (8.24), this often leads to the inequality $\alpha \gg 1$. We shall consider some particular cases in Part III.

We now turn to a consideration of the recombination processes A_{rec} and B_{rec}, i. e. recombinations with emission of quanta and those equivalent to a triple collision. The probability of recombination for the case A_{rec} has already been calculated. It is given by formula (5.50), where the quantity β_ν is related to the absorption coefficient k'_ν by formula (5.55).

The number of recombinations to the kth orbit which take place in 1 second in 1 cm^3 between electrons and $r+1$ times ionised atoms is, in accordance with (5.50) and (5.52),

$$n_e\, n_{r+1}\, 4\pi \left(\frac{m_e}{2\pi kT}\right)^{3/2} \int_0^\infty (\beta_\nu)_k \left(1 + \frac{c^3}{8\pi h\nu^3}\varrho_\nu\right) e^{-m_e v^2/2kT} v^3\, dv\,, \quad (8.25)$$

since all electrons can recombine, whatever their velocities, from 0 to ∞. We have specially added the suffix k to β_ν, in order to show that the electrons are captured into the level k. The total number of photo-recombinations to all levels is

$$Z_{pr} = n_e n_{r+1}\, 4\pi \left(\frac{m_e}{2\pi kT}\right)^{3/2} \sum_k \int_0^\infty (\beta_\nu)_k \left(1 + \frac{c^3}{8\pi h\nu^3 \varrho_\nu}\right) e^{-m_e v^2/2kT} v^3 dv. \quad (8.26)$$

Let us now consider recombination processes which involve triple collisions. Since here the probability of recombination is determined by the simultaneous interaction of three particles, the value of Z_{cr} will be, in general, proportional to $n_e^2\, n_{r+1}$ or to $n_e\, n_H\, n_{r+1}$, according as the third particle is an electron or a hydrogen atom. The part played by metal atoms as third particles is small, since they are generally few in number. However, in all cases the ratio Z_{cr}/Z_{pr} is proportional to the number of particles in 1 cm^3. Since the latter is small (small density of photospheric matter), the ratio in question will be some orders of magnitude less than unity. Contrary to what is the case with ionisation processes, this ratio does not increase markedly when the ratio T_ε/T increases, since the probability of a triple encounter can increase only very slowly with T_ε, and for fairly large T_ε it even decreases when T_ε increases. All this shows that under the conditions in stellar photospheres and even those in the solar corona, where the ratio T_ε/T is of the order of 200, we can usually neglect triple-collision recombinations.

5. **The ionisation equation. The electron temperature.** It follows from the above remarks that the main processes which determine the state of ionisation in photospheres are: (a) photo-ionisation, (b) recombination with emission of quanta (photo-recombination). To construct an equation describing the ionisation state in stellar photospheres, we must equate the quantities Z_{pr} and Z_{pi} from formulae (8.26) and (8.7). This gives

$$\frac{n_{r+1}}{n_r} n_e = \frac{\sum_k \int_{\nu_k}^{\infty} J_\nu \frac{(k_\nu')_k}{\nu} \frac{n_{r,k}}{n_r} \mathrm{d}\nu}{h \left(\frac{m_e}{2\pi kT}\right)^{3/2} \sum_k \int_0^{\infty} (\beta_v)_k \left(1 + \frac{c^2}{8\pi h \nu^3} \varrho_\nu\right) e^{-m_e v^2/2kT} v^3 \mathrm{d}v} \tag{8.27}$$

For the ratio $n_{r,k}/n_r$ we can use formula (5.54). We shall consider in Part II the question of possible deviations from the Boltzmann distribution.

The first question concerning the application of the expression (8.27) is that of the kinetic temperature appearing in the denominator. However, it must be noted at once that the degree of ionisation, as calculated from formula (8.27), depends only slightly on the value taken for T_ε. In the case considered this temperature determines only the probability of photo-recombination, while this probability, as calculations show, changes very slowly with T_ε. Hence we should be interested only in the possibility of large deviations of T_ε from the temperature of the radiation which determines the ionisation. Let us consider this question. Apart from the outermost parts of the photosphere, the value of J_ν in the parts of the spectrum which are the most important for photo-ionisation is close to that of B_ν for the local temperature at the level considered, i. e. for these parts of the spectrum, the density of the ionising radiation is close to the equilibrium density. In other words, the ionisation state of the atoms (which is here the state of *photo-ionisation*) is close to the equilibrium state. However, does this mean that the electron temperature T_ε at every point of the photosphere is equal to the temperature T of the radiation at this point, i. e. the temperature which determines the function $J_\nu = B_\nu$? In order to answer this question, we consider the relation between T and T_ε at the boundary of the photosphere. In accordance with (8.9) we assume that the equality $J_\nu \approx W B_\nu(T_0)$ holds for the most effective ionising frequencies. The case where $J_\nu = B_\nu$ is a particular instance of the more general case, obtained by taking $W = 1$. Next, in order to facilitate the investigation, we restrict ourselves to the first level of the atom, i. e. in the sums appearing in (8.27) we take only the first term, with $\nu_k = \nu_1$. In all practical cases we can neglect the effect of stimulated processes for the frequencies of photo-ionisation from the first level. By virtue of all these assumptions, formula (8.27) takes the form

$$\frac{n_{r+1}}{n_r} n_e = \frac{W \int_{\chi_r/h}^{\infty} B_\nu(T_0) \frac{(k_\nu')_1}{\nu} \mathrm{d}\nu}{h \left(\frac{m_e}{2\pi kT_\varepsilon}\right)^{3/2} \int_0^{\infty} (\beta_v)_1 e^{-m_e v^2/2kT_\varepsilon} v^3 \mathrm{d}v} \tag{8.28}$$

In (8.28) we have, for obvious reasons, taken $n_{r,1}/n_r = 1$. It is also clear that by definition $h\nu_1 = \chi_r$.

We have written equation (8.27), and therefore (8.28), starting from equilibrium between the number of photo-ionisations and the number of photo-recombinations. Such an equilibrium, however, must also exist between the energy absorbed in photo-ionisations and that emitted in photo-recombinations. In this case the defining expressions (8.6) and (8.25) become

$$4\pi \int_{\nu_k}^{\infty} J_\nu (k_\nu')_k \, n_{r,k} \, \mathrm{d}\nu \,, \tag{8.29}$$

$$n_e \, n_{r+1} \, 4\pi \left(\frac{m_e}{2\pi kT}\right)^{3/2} \int_0^{\infty} (\beta_\nu)_k \left(1 + \frac{c^3}{8\pi h \nu^3} \varrho_\nu\right) e^{-m_e v^2/2kT} h\nu \, v^3 \, \mathrm{d}v \,. \tag{8.30}$$

The first of these gives the total energy absorbed in photo-ionisations from the kth level, and the second, the total energy emitted as a result of recombinations to the kth level. The only difference of (8.29) from (8.6) and of (8.30) from (8.25) consists in the appearance of the additional factor $h\nu$ in the integrands.

Returning again to the atom with only one level ($\nu_k = \nu_1$), we can now write down the equality between the corresponding energies. To do so, we must equate the two expressions (8.29) and (8.30), taking into account the above remarks, i. e. putting $J_\nu = W B_\nu(T_0)$, assuming $n_{r,1}/n_r = 1$, and neglecting the stimulated emission. As a result we obtain

$$\frac{n_{r+1}}{n_r} n_e = \frac{W \int\limits_{\chi_r/h}^{\infty} B_\nu(T_0) (k_\nu')_1 \, \mathrm{d}\nu}{h \left(\frac{m_e}{2\pi kT_\varepsilon}\right)^{3/2} \int\limits_0^{\infty} (\beta_\nu)_1 \, e^{-m_e v^2/2kT_\varepsilon} \, \nu \, v^3 \, \mathrm{d}v} \tag{8.31}$$

Now, equating the right-hand sides of (8.28) and (8.31), we obtain the equation required, giving the relation between T_0 and T_ε:

$$\frac{\int\limits_{\chi_r/h}^{\infty} B_\nu(T_0) \frac{(k_\nu')_1}{\nu} \, \mathrm{d}\nu}{\int\limits_{\chi_r/h}^{\infty} B_\nu(T_0) (k_\nu')_1 \, \mathrm{d}\nu} = \frac{\int\limits_0^{\infty} (\beta_\nu)_1 \, e^{-m_e v^2/2kT_\varepsilon} v^3 \, \mathrm{d}v}{\int\limits_0^{\infty} (\beta_\nu)_1 \, e^{-m_e v^2/2kT_\varepsilon} \, \nu \, v^3 \, \mathrm{d}v} \tag{8.32}$$

We now introduce, instead of $B_\nu(T_0)$, the expression for it from (3.11); also, using the formula (5.55) for β_ν, we express the quantity $v\,\mathrm{d}v$ in terms of $\mathrm{d}\nu$ by (5.49), and finally we introduce, instead of $\frac{1}{2} m_e v^2$, the

quantity $h\nu - \chi_r$, in accordance with (5.7). The equation (8.32) then takes the form

$$\frac{\displaystyle\int_{\chi_r/h}^{\infty} \frac{\nu^2 (k_\nu')_1 \, d\nu}{e^{h\nu/kT_0} - 1}}{\displaystyle\int_{\chi_r/h}^{\infty} \frac{\nu^3 (k_\nu')_1 \, d\nu}{e^{h\nu/kT_0} - 1}} = \frac{\displaystyle\int_{\chi_r/h}^{\infty} \nu^2 (k_\nu')_1 \, e^{-h\nu/kT_\varepsilon} \, d\nu}{\displaystyle\int_{\chi_r/h}^{\infty} \nu^3 (k_\nu')_1 \, e^{-h\nu/kT_\varepsilon} \, d\nu} \tag{8.33}$$

We notice immediately that this relation between T_0 and T_ε does not depend on W. Hence this expression will also be valid for the deeper layers of the photosphere.

The solution of (8.33) may be carried out, for instance, for hydrogen-like atoms, for which $(k_\nu')_1$ is given by the expression (5.24). Calculations based on formula (8.33) show that, up to very high temperatures, the two temperatures are practically equal. For very high temperatures it is necessary to take account of the effect of stimulated emission. Similar calculations, taking all the levels of the atom into account (and not only the first), give practically the same result.

In the above-mentioned calculations it is supposed that *recombinations* are the only factor leading to a decrease in the energy of the electron gas. If there is any other fairly efficient source of energy loss (not compensated by a corresponding source of energy gain), for example, the loss of energy in consequence of the excitation or ionisation of atoms by electrons, the temperature T_ε of the latter may be less than the temperature T. In fact, a part of the energy of the electrons will be lost in this case through a process quite unconnected with photo-ionisation and recombination. Preliminary calculations by S. L. Belousov show that for the Sun and for stars of the spectral class A 0 the value of T_ε at the boundary of the photosphere differs somewhat from T_0. However, this cannot noticeably affect the degree of ionisation calculated from formulae (8.27) and (8.28), since, as was shown above, only large differences between T_ε and T_0 are important.

If we now consider the regions of the photosphere where the approximate equality $J_\nu \approx B_\nu$ holds, and assume in accordance with the above that $T \approx T_\varepsilon$, we can evaluate the right-hand side of (8.27) by substituting $J_\nu = B_\nu$ and $T = T_\varepsilon$. To do this, it is necessary to use the Boltzmann distribution (5.54), the relation between β_ν and k_ν' given by the expression (5.55), and the relations (5.49), (5.6) and (5.8). As a result we obtain Saha's formula (5.11). This was to be expected, since all the formulae which we have used correspond to a state of thermodynamic equilibrium. Thus, within the range of validity of our earlier conclusions (for instance, the equality of J_ν and B_ν in the parts of the spectrum

which are the most effective in photo-ionisation) and our assumptions ($T \approx T_\varepsilon$ and the Boltzmann distribution), we may suppose that from a certain depth in the stellar photosphere we can use the ordinary ionisation formula (5.11) for the temperature at the point considered.

For the outer parts of the photosphere (lying above this depth) the approximation $J_\nu \approx B_\nu$ becomes inexact because of the operation of the geometrical effect. In order to study the influence of this effect on the ionisation state, let us again consider the boundary of the atmosphere. As before, we shall assume that the equality $J_\nu = W B_\nu(T_0)$ is here valid for the most effective ionising frequencies. Also, we shall again restrict ourselves, for the sake of simplicity, to the consideration of an atom with one level (the lowest). Then, neglecting the stimulated emission (for reasons explained earlier), we can use formula (8.31). It would be possible also to use formula (8.28), which is derived from the general formula (8.27), but here the calculations are a little more complicated than those relating to (8.31).

Introducing into (8.31) the expressions used in obtaining (8.33) from (8.32), we find

$$\frac{n_{r+1}}{n_r} n_e = \frac{2\, W\, u_{r+1} (2\pi m_e k T_\varepsilon)^{3/2} e^{-\chi_r/kT_\varepsilon}}{g_{r,1} h^3} \frac{\int\limits_{\chi_r/h}^{\infty} \frac{(k_\nu')_1 \nu^3 \, d\nu}{e^{h\nu/kT_0} - 1}}{\int\limits_{\chi_r/h}^{\infty} (k_\nu')_1 \nu^3 e^{-h\nu/kT_\varepsilon} \, d\nu}. \tag{8.34}$$

We now consider the case of hydrogen-like atoms, for which $k_\nu' = c_n \nu^{-3}$ [see (5.24)]. Introducing this expression into (8.34) and effecting the integrations, we obtain

$$\frac{n_{r+1}}{n_r} n_e = W \frac{u_{r+1}}{g_{r,1}} \frac{2 (2\pi m_e)^{3/2} (kT_0)^{3/2}}{h^3} \sqrt{\left(\frac{T_\varepsilon}{T_0}\right)} \log_e (1 - e^{-\chi_r/kT_0})^{-1}. \tag{8.35}$$

Recalling that, in the case we are considering (neglecting stimulated emission) the inequality $e^{-\chi_r/kT_0} \ll 1$ holds (it is valid in general for the majority of practical cases), we can neglect, in the binomial expansion of $(1 - e^{-\chi_r/kT_0})^{-1}$, all terms of higher order than $e^{-\chi_r/kT_0}$. In this case the function considered takes the form $1 + e^{-\chi_r/kT_0}$. Finally, using the expansion

$$\log_e (1 + x) = x - \tfrac{1}{2} x^2 + \tfrac{1}{3} x^3 - \ldots \tag{8.36}$$

and rejecting all terms above the first (for the same reason), we obtain instead of (8.35)

$$\frac{n_{r+1}}{n_r} n_e = W \frac{u_{r+1}}{g_{r,1}} \frac{2 (2\pi m_e)^{3/2} (kT_0)^{3/2}}{h^3} \sqrt{\left(\frac{T_\varepsilon}{T_0}\right)} e^{-\chi_r/kT_0}. \tag{8.37}$$

Thus we obtain Saha's formula (5.11), but with the correction factor $W\sqrt{(T_\varepsilon/T_0)}$ (we may suppose $g_{r,1} \approx u_r$). If we again take $T_0 \approx T_\varepsilon$, then only W remains.

Here it must be borne in mind that in stellar photospheres W is little different from unity. In particular, we can use formula (5.11) without a great error even at the boundary itself, though of course the factor W must be introduced for more exact calculations. The effect of the factor W is important in the case of extended atmospheres and particularly in the case of the planetary nebulae.

Formula (8.37) has been obtained for an atom with only one level. However, more laborious calculations which take the higher levels into account show that the result remains practically the same in principle. Moreover, it is to be remembered that our conclusions are bound up with the application of the equation $J_\nu = WB_\nu$. The effect of an error in T_ε is much less important, as we see from (8.37).

6. The application of Kirchhoff's Law. Let us now consider our last topic, that of the applicability of Kirchhoff's formula (3.10) to stellar photospheres. To do so, we must recall that the processes which mainly determine the radiative power of photospheric material are those of recombination. Free-free transitions play a subordinate part. They are important only in the far infra-red region of the spectrum. Hence, in discussing the applicability of the law (3.10), we can limit ourselves to recombination processes. We consider the recombinations which take place to some level k of the $r+1$ times ionised atom.

The amount of energy emitted by 1 cm³ of matter in 1 second, due to these transitions, in the frequency interval from ν to $\nu + d\nu$ and unit solid angle, is determined by formula (5.57). Using (5.52), (5.49), (5.8) and (5.55), we obtain

$$j_\nu(0)\,\varrho\,d\nu =$$
$$= \{n_e\, n_{r+1}\}\left\{\frac{g_{r,k}\,h^4\,\nu^3\,k_\nu'}{u_{r+1}\,c^2\,(2\pi m_e kT_\varepsilon)^{3/2}}\left[1+\frac{c^2}{2h\nu^3}I_\nu(0)\right]e^{-(h\nu-\chi_{r,k})/kT_\varepsilon}\right\}d\nu. \tag{8.38}$$

We put in this formula $I_\nu(0) = B_\nu(T)$, take $T = T_\varepsilon$, and express the product $n_{r+1}\,n_e$ in terms of $n_{r,k}$ by the use of (5.11), (5.12) and (5.54). Using (5.6), we obtain instead of (8.38)

$$j_\nu(0)\,\varrho\,d\nu = n_{r,k}\,k_\nu'\,B_\nu(T)\,d\nu. \tag{8.39}$$

Thus, by applying to (8.38) the formulae of thermodynamic equilibrium (all for the same temperature), we automatically obtain Kirchhoff's Law (3.10), as was to be expected. Hence the question of the applicability of the law (3.10) to stellar photospheres again reduces to that of the applicability to them of the above-mentioned formulae.

The value of $j_\nu(0)$ in formula (8.38) is determined as the product of the two factors in braces. It might be thought that this product is determined only by the ionisation state of the atoms and by the kinetic temperature. This, however, is not so. In essence, j_ν can be divided into two factors, $\varkappa_\nu$ and B_ν, of which $\varkappa_\nu$ determines the optical thickness of the material in the given direction. Hence, in going from the product $n_e n_{r+1}$ to the number of absorbing atoms, it is necessary to use Boltzmann's Law (5.54), and this we have done in obtaining (8.39) from (8.38). However, we defer the question of the applicability of Boltzmann's formula until Part II. At present we are interested only in ionisation. Here, as we have seen, the situation seems favourable, i. e. the deviations from Saha's formula cannot be large.

The second factor in formula (8.38) is determined by the kinetic temperature T_ε of the electrons, and consequently the main question is whether the temperature T which determines the ionisation differs from T_ε. The agreement between the theory of radiative equilibrium and observation indicates that, in the greater part of the photospheric layers, the equality of T and T_ε holds to a sufficient degree of accuracy. If this is so, then the application of Kirchhoff's Law to stellar photospheres is fully justified. However, the question of the relation between T and T_ε of course requires very careful further study.

The problem of the relation of T to T_ε is particularly important for extended photospheres, where the radiation field J_ν may differ very markedly from B_ν. Finally, it must be borne in mind that, besides the processes considered above which maintain T_ε at a given depth, other processes (for example, those of an electromagnetic character) may also occur in stellar photospheres, as is the case in the upper layers of the solar chromosphere and especially in the solar corona. If in some cases there is a considerable difference between T_ε and the value of T which appears in the formula $J_\nu = WB_\nu(T)$, then the application of Kirchhoff's Law (3.10) is no longer justified. The expression (8.38) must be used directly for the coefficient j_ν, and of course recombinations to all levels possible for the given frequency ν must be taken into account. In this case the relation between the kinetic temperature of the photospheric layers and the intensity of the radiation emerging from them is very complex; we shall consider such an instance in Part III, Chapter 19.

Finally, if there are noticeable deviations from the Boltzmann distribution, this also may make the application of formula (3.10) less justifiable.

PART II.

THE FORMATION OF ABSORPTION LINES IN THE SPECTRA OF STARS

Chapter 9. The mechanism of the formation of absorption lines in stellar atmospheres

1. Basic definitions. The presence of absorption lines in the spectrum of a star shows that the intensity of its continuous radiation is reduced in certain definite frequencies corresponding to transitions of electrons from one level to another.

In order to give a quantitative description of this reduction, we introduce what is called the **residual intensity** r_ν, which is the ratio of the intensity I_ν, at a given frequency ν inside the line, to the intensity of the continous spectrum $I_\nu{}^0$, interpolated to this frequency (Fig. 22a):

$$r_\nu = I_\nu / I_\nu{}^0 \,. \quad (9.1)$$

(r_ν is similarly defined for the flux H_ν.)

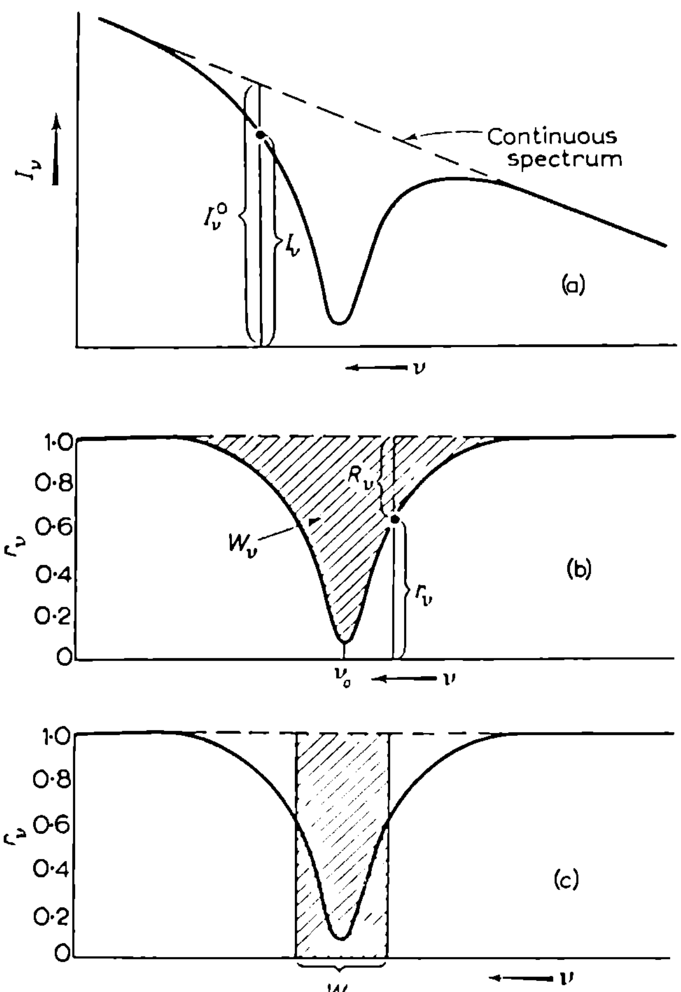

Fig. 22

It is clear that in the frequencies of the continuous spectrum $r_\nu = 1$. As a result we obtain the contour of a line, as shown in Fig. 22 (b). The quantity r_{ν_0}, equal to the least value of r_ν, is called

the **central residual intensity**. The quantity $R_\nu = 1 - r_\nu$ is called the **depth** of the line at a given frequency inside the line. The innermost parts of a line are called its **core**. The outermost parts of a line (in relation to its centre) are called the **wings**.

For many problems, as we shall see later, what is called the **total absorption** in the line, or the **equivalent width of the line**, is of great importance. This quantity is defined by

$$W_\nu = \int (1 - r_\nu)\,\mathrm{d}\nu = \int R_\nu\,\mathrm{d}\nu\,, \tag{9.2}$$

or by

$$W_\lambda = \int (1 - r_\lambda)\,\mathrm{d}\lambda = \int R_\lambda\,\mathrm{d}\lambda\,. \tag{9.3}$$

The symbols ν and λ show that in the first case W is measured on a frequency scale, and in the second case on a wavelength scale. The integration in (9.2) and (9.3) extends over the whole of the line considered. In Fig. 22(b) the quantity W_ν is shown by the shaded area. Observers generally give W on a wavelength scale and measure it in ångströms or in milliångströms. As follows from its definition, the equivalent width is the width of the continuous spectrum whose area is equal to that bounded by the contour of the spectral line considered, i. e. is equal to the shaded area in Fig. 22(b). This is shown in Fig. 22(c).

It is clear that W_λ characterises the strength or intensity of the given line in the stellar spectrum. The greater the equivalent width of an absorption line, the stronger is that line.

2. **The mechanism of the formation of absorption lines.** Let us now investigate how absorption lines are generally formed in stellar spectra. It is evident that at frequencies within absorption lines the absorption coefficient is greater than in the neighbouring frequencies of the continuous spectrum. In the frequencies of an absorption line the absorption coefficient is composed of the coefficient of continuous absorption (free-bound and free-free transitions), added to the absorption coefficient corresponding to the discrete transition in question. This in fact brings about the existence of absorption lines as spectral regions where the photospheric radiation is attenuated by the excess absorption. The fundamental question which arises is how absorption lines are formed: whether in the same way as, for example, the absorption bands are formed beyond the limits of series in the continuous spectra of early-type stars (see Fig. 17), i. e. as a result of *true absorption* processes, or by *scattering of radiation*. In the former case, the attenuation of the radiation in any region of the spectrum is caused by two factors: (1) the absorption coefficient in that region is greater than in the neighbouring regions, (2) the temperature in the photosphere decreases outwards. If the absorption coefficient inside the spectral region considered is very

large, the radiation will emerge from the most superficial and coolest layers of the photosphere; hence the intensity of the radiation emerging from the surface will be close to the intensity $B_\nu(T_0)$, where T_0 is the surface temperature of the star.

Let us now consider the centre of some strong absorption line in the solar spectrum. In this case the absorption coefficient is very large. Hence, for the frequencies of the line centre, the approximate equation $I_\nu \approx B_\nu(T_0)$ holds for any part of the solar disc. On the other hand, according to formula (4.30), the approximate equation $I_\nu \approx B_\nu(T_0)$ holds at the *limb* of the solar disc for all frequencies and, in particular, for the parts of the continuous spectrum adjoining the line considered. Consequently, the value of r_ν for the centre of a strong line should be equal to unity at the limb of the solar disc, i. e. strong lines should be weakened at the limb, and in some cases simply disappear. This is entirely contrary to the observations. At the limb of the solar disc

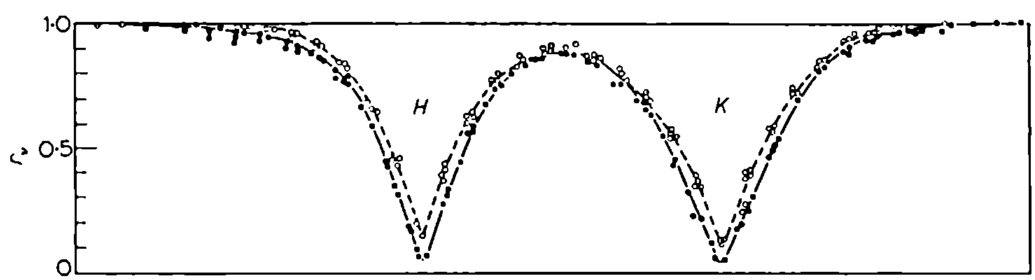

FIG. 23

the contour of a strong line does not differ markedly from the contour of the same line at the centre of the disc. As an illustration we give (Fig. 23) the contours of the H and K lines for two points of the solar disc, the centre (continuous curve) and a point distant from the centre $R = 0{\cdot}98\ R_\odot$ (dotted). (These contours were obtained by V. B. NIKONOV and A. B. SEVERNYĬ.) Consequently, we must reject the first mechanism (true absorption) on observational grounds. This means that, for the description of the processes which determine the formation of absorption lines, Kirchhoff's Law cannot be applied.

In order to illustrate the conclusions we have reached, let us imagine the stellar photosphere to be divided into two layers: a lower, in which the main part of the observed photospheric radiation is produced, and an upper, in which the absorption lines are mainly formed. The lower layer we shall call the photosphere as before, and the upper layer, the reversing layer. Let us assume for simplicity that the reversing layer is composed of atoms which have only two energy levels, the first and the second. Consequently these atoms can absorb and emit only in one fixed line. The excitation of these atoms and the converse transitions from the second state to the first can be brought about

by two causes: (1) by the effect of collisions with the surrounding particles, and (2) as a result of the absorption and subsequent re-emission of light quanta.

If the first cause is the principal one, the Maxwellian distribution of the velocities of the colliding particles leads to the results that the ratio of the number of atoms in the second state to the number in the first state is determined by Boltzmann's Law, and the ratio of the emission coefficient to the absorption coefficient by Kirchhoff's Law (3.10). However, elementary calculations show that, in stellar photospheres, atomic transitions as a result of collisions are very rare in comparison with processes of photo-excitation and subsequent emission. As in the case of ionisation by collisions (see Chapter 8), this is due to the low density of matter in stellar photospheres, i. e. the low concentration of exciting particles. Consequently we must turn to the second cause.

Atoms in the reversing layer absorb the radiation of the photosphere in the frequency considered and scatter this radiation in all directions, including the direction back into the photosphere, where the radiation, after absorption by the atoms of the photosphere, is emitted in other frequencies. The more such atoms there are in the reversing layer, the greater the probability that a quantum of radiation in the frequency considered, having undergone a considerable number of scatterings, falls back into the photosphere and disappears. Thus in this case the part played by the reversing layer amounts to sending back into the photosphere a definite fraction of the quanta which leave the photosphere in the frequency considered: the quanta are diffusely reflected back. In the frequencies of the continuous spectrum, however, the radiation of the photosphere passes through the reversing layer entirely (or almost) unhindered. An absorption line appears as a result of the scattering. Its depth depends on the number of scattering atoms in the reversing layer.

Since the part played by the reversing layer amounts to diffuse selective reflection of radiant energy back into the photosphere, where it is converted into other frequencies, the intensity of radiation in the frequency considered is less in the outer parts of the reversing layer than in the photosphere. The smaller intensity of radiation leads to a smaller proportion of excited atoms. However, the emission coefficient j_ν depends on the number of excited atoms. Hence it is evident that the ratio of j_ν to $\varkappa_\nu$ in the reversing layer must be *less* than in the photosphere, and therefore less than would follow from Kirchhoff's Law.

The assumption that the *chief* mechanism which determines the formation of absorption lines is the *scattering of radiation* is confirmed by observation, and also by a more detailed analysis of the conditions existing in stellar photospheres.

3. The derivation of the equation of transfer for coherent scattering. In this and the four following chapters we shall discuss processes of **coherent** scattering, where the frequency of the re-emitted quantum is *exactly* the same as that of the absorbed quantum, the two quanta differing only in direction. Processes of non-coherent scattering will be considered in Chapter 14.

Assuming now that we are concerned with coherent scattering, let us construct the equation of transfer for these processes. The absorption coefficient corresponding to the discrete transitions which produce the line considered will be denoted by σ_ν; we refer this quantity *to unit mass*. The same coefficient referred *to one atom* will be denoted by s_ν. The relation between them is given by a formula analogous to (5.22).

Let us write down an expression for the emission coefficient in the case of coherent scattering. As is shown by physical analysis, the processes of re-emission of the absorbed quanta are not, in general, *isotropic*. It appears that re-emissions in different directions have different probabilities; in the majority of cases, the probability that the direction of a quantum after scattering will be within the solid angle $d\omega$ is proportional to $(1 + \cos^2 \Phi)\, d\omega$, where Φ is the angle between the incident and scattered quanta. However, an analysis of the problem of scattering, carried out by V. A. AMBARTSUMYAN [9], shows that, for the law mentioned, i. e. $1 + \cos^2 \Phi$, the final results of the theory of scattering do not differ much from those obtained for isotropic scattering. For this reason we shall suppose that the re-emission is *isotropic* in character. An exact allowance for the dependence of the re-emission probability on Φ is one of the most urgent problems of theoretical astrophysics.

Supposing the re-emission to be isotropic, we can immediately write down an expression for j_ν in the case of pure scattering. In fact, according to (3.6) the amount of energy absorbed by one gram of matter in 1 second in unit frequency interval is

$$\sigma_\nu \int_{4\pi} I_\nu \, d\omega \, . \tag{9.4}$$

Now, since we are considering stars *in a steady state* (not varying with time), all this absorbed energy will be restored by the atoms, in the same time interval, in the form of re-emitted quanta. And since we suppose the re-emission to be isotropic, the value of j_ν is obtained by dividing (9.4) by the solid angle of the whole sphere, i. e. 4π. Consequently,

$$j_\nu = \sigma_\nu \int_{4\pi} I_\nu \frac{d\omega}{4\pi} = \sigma_\nu J_\nu \, . \tag{9.5}$$

In writing (9.5) we have obviously assumed that we are concerned with purely coherent scattering, since the absorbed energy is re-emitted in the *same* frequency.

Thus, for isotropic re-emission processes, the difference in the formulae for true absorption and for scattering is that, in the expression for the emission coefficient, the function B_ν is replaced by the mean intensity J_ν. Therefore we can immediately write the equation of transfer for plane-parallel scattering layers in the following form:

$$\cos\theta \, \mathrm{d}I_\nu(\theta)/\mathrm{d}\tau_\nu = I_\nu(\theta) - J_\nu \,, \tag{9.6}$$

where

$$\mathrm{d}\tau_\nu = \sigma_\nu \, \varrho \, \mathrm{d}h \,. \tag{9.7}$$

Here it is clear that the equation of transfer (9.6) takes account only of the change in $I_\nu(\theta)$ due to the presence of coherent scattering processes.

We now multiply both sides of (9.6) by $1/4\pi$ and integrate the result over the whole sphere:

$$\frac{\mathrm{d}}{\mathrm{d}\tau_\nu} \int_{4\pi} I_\nu(\theta) \cos\theta \frac{\mathrm{d}\omega}{4\pi} = \frac{1}{4\pi} \frac{\mathrm{d}(\pi H_\nu)}{\mathrm{d}\tau_\nu} = \int I_\nu(\theta) \frac{\mathrm{d}\omega}{4\pi} - J_\nu \equiv 0 \,. \tag{9.8}$$

Thus, in a purely scattering atmosphere, composed of plane-parallel layers, the flux πH_ν is *constant at all depths*. In contrast to the case of local thermodynamic equilibrium, this constancy of the flux holds, in the case of coherent scattering, for each single frequency.

4. The solution of the equations of transfer. Let us now consider the following model. Above the layers which radiate the continuous spectrum (the "photosphere" of the star) lies a scattering layer (the "reversing layer"). Here we again sharply separate the layers which form the continuous spectrum and those which form the absorption lines. We shall discuss later to what extent this model can represent the actual envelope of a star.

Our main problem is to calculate the flux πH_ν emerging from the surface of the scattering layer. We denote by $I_\nu^0(\theta)$ the intensity of the continuous radiation emerging from the "photosphere". Bearing in mind the difference between the model considered and the actual photosphere, it will be sufficient, in considering the model, to use the same approximate method, introducing mean intensities, as we discussed in Chapter 3.

Denoting again by I_ν and I'_ν the mean intensities of radiation in the appropriate hemispheres and noting that equation (9.6) is similar to equation (3.12), or for the lower hemisphere (introducing the angle $\psi = \pi - \theta$) to equation (3.17), we can immediately write for I_ν and I'_ν, from (3.61) and (3.62),

$$\tfrac{2}{3} \, \mathrm{d}I_\nu/\mathrm{d}\tau_\nu = I_\nu - J_\nu \,, \tag{9.9}$$

$$\tfrac{2}{3} \, \mathrm{d}I'_\nu/\mathrm{d}\tau_\nu = -I_\nu' + J_\nu \,. \tag{9.10}$$

Adding equations (9.9) and (9.10), we obtain

$$\mathrm{d}(I_\nu + I_\nu')/\mathrm{d}\tau_\nu = \tfrac{3}{2}(I_\nu - I_\nu') . \tag{9.11}$$

Noticing that by (3.50) the difference of the mean intensities is H_ν, while, according to what was said above, the value of H_ν in scattering plane-parallel layers is constant, we find

$$I_\nu + I_\nu' = \tfrac{3}{2} H_\nu \tau_\nu + D_\nu , \tag{9.12}$$

where D_ν is a constant of integration. It follows from (9.12) that D_ν is equal to the intensity I_ν at the outer boundary of the reversing layer, where $\tau_\nu = 0$ and $I_\nu' = 0$. In turn, for $\tau_\nu = 0$ (where $I_\nu' \equiv 0$) this intensity is equal to H_ν, in virtue of (3.50). Consequently $D_\nu = H_\nu$, and hence

$$I_\nu + I_\nu' = \tfrac{3}{2} H_\nu \tau_\nu + H_\nu . \tag{9.13}$$

We now determine the value of H_ν. To do so, we write the equations (3.50) and (9.13) for the lower surface of the reversing layer. Denoting the optical thickness of the *whole* reversing layer by $\tau_{\nu,\sigma}$, we have

$$\left.\begin{aligned} (I_\nu)_{\tau_{\nu,\sigma}} - (I_\nu')_{\tau_{\nu,\sigma}} &= H_\nu , \\ (I_\nu)_{\tau_{\nu,\sigma}} + (I_\nu')_{\tau_{\nu,\sigma}} &= \tfrac{3}{2} H_\nu \tau_{\nu,\sigma} + H_\nu . \end{aligned}\right\} \tag{9.14}$$

Adding these two expressions, we obtain

$$H_\nu/(I_\nu)_{\tau_{\nu,\sigma}} = 1/(1 + \tfrac{3}{4}\tau_{\nu,\sigma}) . \tag{9.15}$$

The value of $(I_\nu)_{\tau_{\nu,\sigma}}$ is the mean intensity of radiation at the lower surface of the scattering layer. It is clear, however, that this mean intensity, multiplied by π, is also equal to the flux of photospheric continuous radiation in the frequency considered, i. e. to

$$\pi H_\nu^0 = 2\pi \int_0^{\frac{1}{2}\pi} I_\nu^0(\theta) \cos\theta \sin\theta \, \mathrm{d}\theta . \tag{9.16}$$

In fact, putting $\tau_{\nu,\sigma} \equiv 0$, i. e. supposing that there is no attenuation of the radiation by scattering in the given frequency, we obtain from (9.15) $(I_\nu)_{\tau_{\nu,\sigma}} = H_\nu^0$. The value of πH_ν^0 is just the flux of continuous radiation interpolated to the given frequency. And since the flux πH_ν is the flux in the line, we can write for the residual intensity

$$r_\nu = H_\nu/H_\nu^0 = 1/(1 + \tfrac{3}{4}\tau_{\nu,\sigma}) . \tag{9.17}$$

The residual intensity r_ν determined by the expression (9.17) is derived for the flux of radiation. Consequently, the formula (9.17) is applicable to the radiation of *a distant star* or to the radiation of *the Sun's disc as a whole.* According to this formula, when s_ν (and consequently $\tau_{\nu,\sigma}$) increases, r_ν decreases. In the centre of the line, where s_ν and $\tau_{\nu,\sigma}$ have their greatest value, r_ν is least. Thus we obtain a contour of the form shown in Fig. 22(b).

In order to obtain the contour of an absorption line, not for the whole disc of the Sun, but for a definite point of it, i. e. for some direction θ, we must solve equation (9.6). Starting from the formal identity, already pointed out, between this equation and equation (3.12), we can use the solutions given by formulae (3.26) and (3.27). Here it is necessary to take the latter of these formulae, since it gives the intensity determined by a layer of finite optical thickness. In the case considered we are interested in radiation bounded by the directions $0 \leq \theta \leq \frac{1}{2}\pi$, and hence we must introduce $\theta = \pi - \psi$ into (3.27) instead of ψ. We then obtain for $I_\nu(\theta, \tau_\nu)$

$$I_\nu(\theta, \tau_\nu) = G_\nu\, e^{\tau_\nu \sec\theta} - e^{\tau_\nu \sec\theta} \int\limits_0^{\tau_\nu} J_\nu\, e^{-t_\nu \sec\theta} \sec\theta\, dt_\nu\,, \qquad (9.18)$$

where G_ν is a constant of integration.

For J_ν we can take the expression which follows from (3.51) and (9.13):

$$J_\nu = \tfrac{1}{2} H_\nu (1 + \tfrac{3}{2} t_\nu)\,. \qquad (9.19)$$

According to (9.18), we have at the *upper* (outer) boundary of the reversing layer, where $\tau_\nu = 0$, the equation

$$I_\nu(\theta, 0) = G_\nu\,. \qquad (9.20)$$

Also, at the *lower* boundary of the reversing layer, where $\tau_\nu = \tau_{\nu,\sigma}$, we have

$$I_\nu(\theta, \tau_{\nu,\sigma}) = I_\nu^0(\theta)\,. \qquad (9.21)$$

Writing (9.18) for this level and solving for G_ν the resulting expression, we obtain by (9.20)

$$I_\nu(\theta,0) = I_\nu^0(\theta)\, e^{-\tau_{\nu,\sigma} \sec\theta} + \int\limits_0^{\tau_{\nu,\sigma}} J_\nu\, e^{-t_\nu \sec\theta} \sec\theta\, dt_\nu\,. \qquad (9.22)$$

Finally, substituting in (9.22) the expression (9.19) and effecting the integration, we find

$$I_\nu(\theta,0) = I_\nu^0(\theta)\, e^{-\tau_{\nu,\sigma}\sec\theta} + \\ + \frac{H_\nu}{2}\left\{ (1 + \tfrac{3}{2}\cos\theta) - e^{-\tau_{\nu,\sigma}\sec\theta}(1 + \tfrac{3}{2}\tau_{\nu,\sigma} + \tfrac{3}{2}\cos\theta)\right\}, \quad (9.23)$$

where we may substitute for H_ν the expression for it given by (9.17).

The residual intensity $r_\nu(\theta)$ in the case considered must be determined as the ratio $I_\nu(\theta,0)/I_\nu^0(\theta)$. Using the definition of the flux H_ν^0, it is

$$r_\nu(\theta) = I_\nu(\theta,0)/I_\nu^0(\theta)$$

$$= e^{-\tau_{\nu,\sigma}\sec\theta} + \frac{1}{1+\frac{3}{4}\tau_{\nu,\sigma}} \frac{\int\limits_0^{\frac{1}{2}\pi} I_\nu^0(\theta)\cos\theta\sin\theta\,d\theta}{I_\nu^0(\theta)} \Big\{(1+\tfrac{3}{2}\cos\theta) -$$

$$- e^{-\tau_{\nu,\sigma}\sec\theta}(1+\tfrac{3}{2}\tau_{\nu,\sigma}+\tfrac{3}{2}\cos\theta)\Big\}. \qquad (9.24)$$

Consequently, in order to find the dependence of $r_\nu(\theta)$ on the frequency for a given angle θ, we must know: (1) the law of variation of $\tau_{\nu,\sigma}$ with frequency, (2) the law of darkening of the disc towards the limb for the intensity of the continuous radiation in the given spectral region, i. e. the variation of $I_\nu^0(\theta)$ as a function of the angle θ. The latter may be taken directly from observation.

We shall discuss later, when considering the form of the dependence of σ_ν on the frequency, the comparison of formula (9.24), as well as of formula (9.17), with observation. In particular, we shall see that, for the wings of some lines in solar and stellar spectra, these formulae lead to a satisfactory agreement of the results of theory and observation, this being achieved by an appropriate choice of the parameters appearing in $\tau_{\nu,\sigma}$.

Formula (9.24) shows, in agreement with observation, that the variation of absorption-line contours from the centre to the limb of the solar disc is generally small*.

However, the agreement between the above theory and observation is very often not at all satisfactory. A quantitative comparison between the results of the theory we have developed and observation discloses a number of discrepancies. Thus, for instance, in the centres of strong absorption lines, where $\tau_{\nu,\sigma}$ is very large, the central residual intensities ought to be close to zero, from formulae (9.17) and (9.24). However, such a result is contrary to observation. This, of course, is explained by the fact that our model does not exactly correspond to actual photospheres.

---

* For the centres of strong lines, where $\tau_{\nu,\sigma}$ is very large, this follows immediately from (9.24), since, on the one hand, for $\tau_{\nu,\sigma} \to \infty$ the residual intensity $r_\nu(\theta)$ is proportional to the ratio $(1+\frac{3}{2}\cos\theta)/I_\nu^0(\theta)$, while, on the other hand, in the visible part of the spectrum the intensity $I_\nu^0(\theta)$ itself approximately follows the law $(1+\frac{3}{2}\cos\theta)$, so that in general the variation of $r_\nu(\theta)$ with the angle θ is small.

5. **The derivation of the equation of transfer, taking account of processes of scattering and of true absorption.** The main simplifying assumption in the model we have analysed is the artificial division of the atmosphere of the star into two quite different regions, the "photosphere" and the "reversing layer". In reality, of course, no such division exists. In fact, the power of scattering is possessed not only by the atoms of the reversing layer, but by those of the photosphere also. Thus, in the formation of an absorption line, *all* the photospheric layers of the star (or the Sun) take part. It follows from this that, for a rigorous solution of the problem in question, we must abandon the division of the atmosphere into the photosphere and the reversing layer, and start from the *general* equation of transfer, which takes account not only of scattering processes, but also of those of absorption.

To construct this equation, we must start from the fact that, in a real atmosphere, the attenuation of a beam of radiation in a segment $\mathrm{d}s$ is determined, in the frequencies of an absorption line, not only by the selective absorption coefficient σ_ν, but also by the continuous absorption coefficient $\varkappa_\nu$. Thus the amount of radiation absorbed in a cylinder of volume $\mathrm{d}\sigma\,\mathrm{d}s$ (see Fig. 3), in the frequency interval $\mathrm{d}\nu$, solid-angle element $\mathrm{d}\omega$ and time element $\mathrm{d}t$, is, by (2.25),

$$\Delta E_{\mathrm{ab}} = -\, I_\nu (\varkappa_\nu + \sigma_\nu)\, \varrho\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}\sigma\, \mathrm{d}t\, \mathrm{d}\nu\,. \tag{9.25}$$

We now write down the amount of energy emitted by the cylinder as a result of processes of *true continuous absorption*. In accordance with (2.21) and (3.10), we can write for these processes

$$\Delta E_{\mathrm{em}}{}^{(1)} = \varkappa_\nu\, B_\nu\, \varrho\, \mathrm{d}\sigma\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}t\, \mathrm{d}\nu\;, \tag{9.26}$$

since for processes of true absorption we assume the laws of thermodynamic equilibrium to be applicable.

The processes of the selective attenuation of radiation, i. e. those taking place only at frequencies within lines, require special attention. These processes include: (1) scattering proper of radiation, which has already been considered in the present chapter, (2) true selective absorption, caused by discrete transitions of electrons. These latter processes were discussed in Section 5.3. Let us first calculate ΔE_{em} for scattering processes. Of the total energy *selectively* absorbed by the cylinder, i. e. of the amount $I_\nu\, \sigma_\nu\, \varrho\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}\sigma\, \mathrm{d}t\, \mathrm{d}\nu$, let the fraction $1 - \varepsilon_\nu$ be re-emitted as *coherently scattered radiation*, so that the fraction ε_ν of this energy will be expended on processes of true selective absorption. In other words, the quantity of absorbed energy which is re-emitted in the form of scattered radiation is $I_\nu\, \sigma_\nu (1 - \varepsilon_\nu) \varrho\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}\sigma\, \mathrm{d}t\, \mathrm{d}\nu$. This means that the absorption coefficient wich corresponds only to pro-

cesses of subsequent scattering is not σ_ν, but $\sigma_\nu(1-\varepsilon_\nu)$. Consequently, in accordance with (9.5) and (2.21), the amount of energy re-emitted by the cylinder as a result of coherent scattering processes is

$$\Delta E_{\mathrm{em}}^{(2)} = \sigma_\nu (1-\varepsilon_\nu)\, J_\nu\, \varrho\, \mathrm{d}\sigma\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}t\, \mathrm{d}\nu\, . \tag{9.27}$$

We now turn to the radiation which corresponds to processes of true selective absorption. The amount of energy, absorbed (from the beam) in the cylinder, which corresponds to these processes is, by what has been said above, $I_\nu\, \sigma_\nu\, \varepsilon_\nu\, \varrho\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}\sigma\, \mathrm{d}t\, \mathrm{d}\nu$; consequently the absorption coefficient belonging to true selective absorption is $\sigma_\nu\, \varepsilon_\nu$. And since, for processes of true absorption, we assume Kirchhoff's Law to be applicable (i. e. we assume the existence of local thermodynamic equilibrium), the amount of energy emitted by the cylinder which corresponds to processes of true selective absorption is, by (2.21) and (3.10),

$$\Delta E_{\mathrm{em}}^{(3)} = \sigma_\nu\, \varepsilon_\nu\, B_\nu\, \varrho\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}t\, \mathrm{d}\nu\, \mathrm{d}\sigma\, . \tag{9.28}$$

We notice the following difference between processes of scattering and those of true selective absorption: in scattering, *all* the absorbed energy $I_\nu\, \sigma_\nu(1-\varepsilon_\nu)\, \varrho\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}\sigma\, \mathrm{d}t\, \mathrm{d}\nu$ is again re-emitted in the form of scattered energy, while for processes of true absorption this equality between the absorbed and re-emitted energy does not hold. The amount of emitted energy, expressed by formula (9.28), has no direct relation to the amount of absorbed energy, i. e. $I_\nu\, \sigma_\nu\, \varepsilon_\nu\, \varrho\, \mathrm{d}s\, \mathrm{d}\omega\, \mathrm{d}\sigma\, \mathrm{d}t\, \mathrm{d}\nu$. This corresponds to the statement made in Section 5.3, that in the presence of true absorption processes there is no direct relation between the absorbed and emitted quanta.

Now, constructing the balance between absorbed and emitted energy as in the derivation of (2.26), and using (2.27), we obtain the required equation of transfer:

$$\cos\theta\, \frac{\mathrm{d}I_\nu(\theta)}{\varrho\, \mathrm{d}h} = (\varkappa_\nu + \sigma_\nu)\, I_\nu(\theta) - (1-\varepsilon_\nu)\, \sigma_\nu\, J_\nu - \varkappa_\nu\, B_\nu - \sigma_\nu\, \varepsilon_\nu\, B_\nu\, . \tag{9.29}$$

If processes of true selective absorption are disregarded ($\varepsilon_\nu = 0$), this equation takes the form

$$\cos\theta\, \frac{\mathrm{d}I_\nu(\theta)}{\varrho\, \mathrm{d}h} = (\varkappa_\nu + \sigma_\nu)\, I_\nu(\theta) - \sigma_\nu\, J_\nu - \varkappa_\nu\, B_\nu\, . \tag{9.30}$$

The equation (9.29) can be modified by putting

$$(1-\varepsilon_\nu)\, \sigma_\nu = \sigma_\nu^0\, , \quad \sigma_\nu\, \varepsilon_\nu = \varkappa_\nu^0\, . \tag{9.31}$$

It then takes the form

$$\cos\theta\, \frac{\mathrm{d}I_\nu(\theta)}{\varrho\, \mathrm{d}h} = (\varkappa_\nu + \sigma_\nu^0 + \varkappa_\nu^0)\, I_\nu(\theta) - \sigma_\nu^0\, J_\nu - (\varkappa_\nu + \varkappa_\nu^0)\, \mathrm{B}_\nu\, . \tag{9.32}$$

The physical meaning of the coefficients σ_ν^0 and $\varkappa_\nu^0$ which have been introduced is quite clear. We may call the coefficient σ_ν^0 the **scattering coefficient in the line**, since it takes account only of scattering processes (not including those of true absorption). The coefficient $\varkappa_\nu^0$ is the **coefficient of true selective absorption.**

In the next chapter we consider the mathematical methods by which the above-mentioned equations are solved.

Chapter 10. The solution of the equations of transfer for frequencies inside absorption lines

1. **Introductory remarks.** The problem of solving the equations of transfer (9.29), (9.30), and (9.32) is extremely complicated. Firstly, all the physical parameters appearing in these equations, i. e. $I_\nu(\theta)$, $\varkappa_\nu$, σ_ν, ε_ν, B_ν, and ϱ, vary with depth, and the dependence on h is different for each of them. Secondly, difficulties arise concerning the dependence of the intensity $I_\nu(\theta)$ on the angle θ. Hence it appears impossible to obtain at once a general solution of the equations in question. It is therefore necessary to introduce various simplifying assumptions. We first consider the solution of the equation of transfer under the assumption that the ratio of the absorption coefficient σ_ν in the line to the continuous absorption coefficient $\varkappa_\nu$ is the same, for a given ν, at all depths. This assumption is fairly well satisfied for the centres of strong absorption lines.

A second method is the numerical integration of the equation of transfer. This method leads to exact results, but is very laborious.

Finally, we shall analyse the solution of the equation of transfer under the assumption that both the scattering coefficient in the line σ_ν^0 and the true selective absorption coefficient $\varkappa_\nu^0$ are very small compared with the continuous absorption coefficient $\varkappa_\nu$. The solution for this special case is of great importance for faint absorption lines and for the wings of medium and strong lines.

2. **A model of the atmosphere with a constant ratio of the absorption coefficient in the line to the continuous absorption coefficient.** We introduce the following simplifying assumptions: (1) we shall suppose that throughout the atmosphere of the star the ratio

$$\eta_\nu = \sigma_\nu/\varkappa_\nu \tag{10.1}$$

is constant for any given frequency; (2) the same assumption is made regarding the value of ε_ν; (3) we shall suppose that the linear expansion (8.10) of the Planck function B_ν is applicable. The optical depth t_ν is here defined in terms of the continuous absorption coefficient $\varkappa_\nu$.

Dividing (9.29) by $\varkappa_\nu$ and using (10.1), we obtain

$$\cos\theta\, \mathrm{d}I_\nu(\theta)/\mathrm{d}\tau_\nu = (1+\eta_\nu)\, I_\nu(\theta) - (1-\varepsilon_\nu)\, \eta_\nu J_\nu - B_\nu - \eta_\nu\, \varepsilon_\nu\, B_\nu\,, \quad (10.2)$$

where

$$\mathrm{d}\tau_\nu = \varkappa_\nu\, \varrho\, \mathrm{d}h\,. \quad (10.3)$$

We multiply both sides of equation (10.2), first by $\cos\theta/4\,\pi$ and then by $1/4\,\pi$, and in both cases integrate the result over the whole sphere. Taking account of the definitions (2.5), (3.15) and (3.46), we obtain respectively

$$\mathrm{d}K_\nu/\mathrm{d}\tau_\nu = \tfrac{1}{4}\,(1+\eta_\nu)\, H_\nu\,, \quad (10.4)$$

$$\tfrac{1}{4}\, \mathrm{d}H_\nu/\mathrm{d}\tau_\nu = (1+\varepsilon_\nu\, \eta_\nu)\, (J_\nu - B_\nu)\,. \quad (10.5)$$

We introduce the first mathematical simplification by assuming that the equation (3.54) holds also in the case we are considering. [This equation is rigorously true if $I_\nu(\theta)$ can be represented as the sum of a constant term and some odd function of $\cos\theta$.] Assuming the validity of (3.54), we can rewrite equation (10.4) as

$$\mathrm{d}J_\nu/\mathrm{d}\tau_\nu = \tfrac{3}{4}\,(1+\eta_\nu)\, H_\nu\,. \quad (10.6)$$

We differentiate (10.6) with respect to τ_ν, taking into account the fact that, by our assumption, η_ν = constant throughout the atmosphere of the star:

$$\mathrm{d}^2 J_\nu/\mathrm{d}\tau_\nu = \tfrac{3}{4}\,(1+\eta_\nu)\, \mathrm{d}H_\nu/\mathrm{d}\tau_\nu\,. \quad (10.7)$$

Next, introducing in (10.7) the expression for the derivative $\mathrm{d}H_\nu/\mathrm{d}\tau_\nu$ from (10.5), we obtain

$$\mathrm{d}^2 J_\nu/\mathrm{d}\tau_\nu^2 = q_\nu^2\, (J_\nu - B_\nu)\,, \quad (10.8)$$

where

$$q_\nu = \sqrt{[3(1+\eta_\nu)\,(1+\varepsilon_\nu\,\eta_\nu)]}\,. \quad (10.9)$$

According to our assumption, B_ν is a linear function of optical depth, and hence the second derivative of B_ν with respect to τ_ν is zero. This makes it possible to rewrite (10.8) in the following form:

$$\mathrm{d}^2(J_\nu - B_\nu)/\mathrm{d}\tau_\nu^2 = q_\nu^2\,(J_\nu - B_\nu)\,. \quad (10.10)$$

The solution of equation (10.10) can be represented in the following form:

$$J_\nu - B_\nu = C_\nu\, e^{-q_\nu \tau_\nu}\,. \quad (10.11)$$

We have not here added a term of the form $D_\nu e^{q_\nu \tau_\nu}$ to the right-hand side of (10.11), since, according to the results of Chapter 6, the value of J_ν as $\tau_\nu \to \infty$ must tend to B_ν, and consequently $D_\nu \equiv 0$.

Let us determine the constant C_ν. From (10.11) and (8.10) we can write

$$J_\nu = a_\nu + b_\nu \tau_\nu + C_\nu e^{-q_\nu \tau_\nu} . \tag{10.12}$$

Differentiating this expression with respect to τ_ν and substituting the result into (10.6), we obtain for H_ν

$$H_\nu = \tfrac{4}{3} \frac{1}{1 + \eta_\nu} \{b_\nu - C_\nu q_\nu e^{-q_\nu \tau_\nu}\} . \tag{10.13}$$

For the boundary of the stellar atmosphere, where $\tau_\nu = 0$, the flux $H_\nu(0)$ is

$$H_\nu(0) = \tfrac{4}{3} \frac{1}{1 + \eta_\nu} \{b_\nu - C_\nu q_\nu\} . \tag{10.14}$$

On the other hand, the value of J_ν for $\tau_\nu = 0$ is, according to (10.12),

$$J_\nu(0) = a_\nu + C_\nu . \tag{10.15}$$

As a boundary condition we shall take the following:

$$2J_\nu(0) = H_\nu(0) . \tag{10.16}$$

This equation, which is of course approximate, follows from (3.50) and (3.51), on taking into account the fact that $I_\nu' \equiv 0$ at the boundary of the atmosphere. Introducing now the expressions (10.14) and (10.15) into (10.16), and solving the equation obtained for the constant C_ν, we find

$$C_\nu = \frac{\tfrac{2}{3} b_\nu - a_\nu (1 + \eta_\nu)}{\tfrac{2}{3} q_\nu + (1 + \eta_\nu)} . \tag{10.17}$$

Substituting this expression for C_ν in (10.14), we obtain

$$H_\nu(0) = \tfrac{4}{3} \frac{b_\nu + a_\nu q_\nu}{1 + \eta_\nu + \tfrac{2}{3} q_\nu} . \tag{10.18}$$

This is the "*observed*" *flux* of stellar radiation for a given frequency inside the line, which is what interests us. The flux H_ν^0, corresponding to the continuous spectrum interpolated to this frequency, is found by putting $\eta_\nu = 0$. This means that there is no selective attenuation of the radiation in the frequencies of the continuous spectrum. From (10.18) and (10.9), the value of $H_\nu^0(0)$ is

$$H_\nu^0(0) = \frac{4 (b_\nu + a_\nu \sqrt{3})}{3 (1 + \tfrac{2}{3} \sqrt{3})} . \tag{10.19}$$

Hence we have for the residual intensity r_ν

$$r_\nu = \frac{H_\nu(0)}{H_\nu^0(0)} = \frac{b_\nu + a_\nu q_\nu}{b_\nu + a_\nu \sqrt{3}} \frac{1 + \frac{2}{3}\sqrt{3}}{1 + \eta_\nu + \frac{2}{3} q_\nu}. \tag{10.20}$$

We notice that, in deriving this expression for r_ν, we have used two approximate equations: (1) $J_\nu = 3K_\nu$, (2) $H_\nu(0) = 2J_\nu(0)$. Both these equations are borrowed from the method using averaged intensities which we considered in Chapter 3. A more exact discussion of the problem leads to the following results: (1) the approximation $J_\nu = 3K_\nu$ is sufficiently exact for practical purposes; (2) the approximation $H_\nu(0) = 1{\cdot}8\, J_\nu(0)$ is more exact than $H_\nu(0) = 2J_\nu(0)$.

In order to find r_ν from formula (10.20) for any frequency inside the line, it is necessary to know the values of η_ν and ε_ν for this frequency, and also the coefficients a_ν and b_ν. In the general case both η_ν and ε_ν vary with depth, sometimes very markedly. Moreover, in general the function B_ν is not a linear function of τ_ν. The assumption of its linearity often leads to serious errors in the final result.

The case considered above is not far from the truth for the central parts of medium and strong absorption lines. For the wings of lines this model, strictly speaking, can be applied only in particular cases.

The coefficients a_ν and b_ν must be so chosen that the linear approximation to B_ν best represents the actual law of variation of B_ν with τ_ν. This law is found from the theory of stellar photospheres.

We shall discuss the actual application of formula (10.20) in the following chapters.

Since formula (10.12) together with (10.17) establishes the dependence of J_ν on τ_ν, we can solve equation (10.2) not only for $H_\nu(0)$, but also for $I_\nu(0,0)$. This is done in the same way as the solution of equations (3.12) and (3.17).

3. **The numerical integration of the equations of transfer.** The chief defect of the solution, considered above, of equation (10.2) is the introduction of simplifying assumptions concerning the functions η_ν, ε_ν and B_ν. The simplifications which are of a purely mathematical character [$J_\nu = 3K_\nu$ and $2J_\nu(0) = H_\nu(0)$] as a rule introduce much smaller errors. Bearing this in mind, we can solve equation (10.2) numerically, starting from the two equations (10.5) and (10.6).

Let us consider a method of numerical integration of equation (10.2). The dependence of B_ν on τ_ν, for a given frequency, is derived from the theory of photospheres, which gives, for each level in the photosphere, the temperature T, the pressures p and p_e, the density ϱ, the optical depth τ_ν (τ_ν is here defined in terms of the continuous absorption coefficient), etc. From these parameters and the appropriate expressions for

$\varkappa_\nu$ and σ_ν, we can determine the value of η_ν, and independently that of ε_ν, for every level.

As boundary conditions for the integration of equations (10.5) and (10.6) it is most convenient to take the following: (1) $2J_\nu(0) = H_\nu(0)$, (2) as $\tau_\nu \to \infty$, the value of J_ν tends to B_ν. The latter condition is evident from the results of Chapter 6. The factor 2 in the former condition can be replaced, as we have said, by the more accurate factor 1·8.

The integration of equations (10.5) and (10.6) is performed as follows. Having taken various definite values of $J_\nu(0)$ [and consequently of $H_\nu(0)$, from the condition $2J_\nu(0) = H_\nu(0)$], we carry out the numerical integration of equations (10.5) and (10.6) for each of these values. The value $J_\nu{}^*(0)$ for which J_ν tends to B_ν for large τ_ν is *the required boundary value*. The flux $\pi H_\nu(0)$ corresponding to this boundary value $J_\nu{}^*(0)$ is the solution of the problem proposed. Since the variation of J_ν with τ_ν is found by this method, we can also find, by means of the original equation, the emergent intensity $I_\nu(\theta, 0)$ at the surface of the Sun, for various angles θ.

The numerical method described is generally sufficiently accurate, the more so since from the results obtained we can estimate a more reliable ratio of $J_\nu(0)$ to $H_\nu(0)$ and thereby again carry out the integration of equations (10.5) and (10.6). However, this method is inconvenient, since it requires a great expenditure of time on the numerical integration. In fact, the integration of equations (10.5) and (10.6) has to be performed for many points (frequencies) inside the line. Hence a more rapid method is to be preferred.

4. The solution of the equation of transfer for faint lines. As a third method, we shall discuss the solution of equation (9.32) given by A. Unsöld for faint lines and for the outermost wings of strong and medium lines. This method is unsuitable for residual intensities $r_\nu < 0{\cdot}85$. As we shall see, faint lines play an extremely important part in a number of cases, and particularly in the problem of the chemical composition of the photosphere. Hence the theory mentioned deserves consideration.

In order to solve equation (9.32), we introduce the function S_ν, defined by the equation

$$\sigma_\nu{}^0 J_\nu + (\varkappa_\nu + \varkappa_\nu{}^0) B_\nu = (\varkappa_\nu + \sigma_\nu{}^0 + \varkappa_\nu{}^0) S_\nu \,. \tag{10.21}$$

The equation (9.32) then takes the form

$$\cos\theta \frac{\mathrm{d}I_\nu(\theta)}{\varrho\,\mathrm{d}h} = (\varkappa_\nu + \sigma_\nu{}^0 + \varkappa_\nu{}^0) I_\nu(\theta) - (\varkappa_\nu + \sigma_\nu{}^0 + \varkappa_\nu{}^0) S_\nu \,. \tag{10.22}$$

We now introduce the optical depth $\mathrm{d}z_\nu (= \mathrm{d}x_\nu)$:

$$\mathrm{d}z_\nu = (\varkappa_\nu + \sigma_\nu{}^0 + \varkappa_\nu{}^0)\,\varrho\,\mathrm{d}h \,. \tag{10.23}$$

Then (10.22) may be written

$$\cos\theta \, \mathrm{d}I_\nu(\theta)/\mathrm{d}z_\nu = I_\nu(\theta) - S_\nu(z_\nu) \, . \tag{10.24}$$

We shall now obtain an integral equation for $S_\nu(z_\nu)$.

The equation of transfer (10.24) is identical with the fundamental equation (3.12). Hence we can use the solutions (3.36) and (3.37) for the intensities in the respective hemispheres, replacing B_ν by the function S_ν. As a result, we obtain for J_ν

$$\begin{aligned} J_\nu &= \int I_\nu(\theta) \, \mathrm{d}\omega/4\pi \\ &= \tfrac{1}{2} \{ \int_0^{\frac{1}{2}\pi} I_\nu(\theta, z_\nu) \sin\theta \, \mathrm{d}\theta + \int_0^{\frac{1}{2}\pi} I_\nu'(\psi, z_\nu) \sin\psi \, \mathrm{d}\psi \} \\ &= \tfrac{1}{2} \{ \int_0^{\frac{1}{2}\pi} [\int_{z_\nu}^{\infty} S_\nu(x_\nu) \, e^{-(x_\nu - z_\nu)\sec\theta} \sec\theta \, \mathrm{d}x_\nu] \sin\theta \, \mathrm{d}\theta + \\ &+ \int_0^{\frac{1}{2}\pi} [\int_0^{z_\nu} S_\nu(x_\nu) \, e^{-(z_\nu - x_\nu)\sec\psi} \sec\psi \, \mathrm{d}x_\nu] \sin\psi \, \mathrm{d}\psi \} \, . \end{aligned} \tag{10.25}$$

We change the order of integration in (10.25) and introduce, instead of θ and ψ, a new variable w by means of the relations $\sec\theta = w$ and $\sec\psi = w$. Then we can rewrite (10.25) as

$$J_\nu = \tfrac{1}{2} \{ \int_{z_\nu}^{\infty} S_\nu(x_\nu) \, E_1(x_\nu - z_\nu) \, \mathrm{d}x_\nu + \int_0^{z_\nu} S_\nu(x_\nu) \, E_1(z_\nu - x_\nu) \, \mathrm{d}x_\nu \} \, , \tag{10.26}$$

where the function E_1 is defined by the expression (6.35). We can write the expression (10.26) as

$$J_\nu(z_\nu) = \tfrac{1}{2} \int_0^{\infty} S_\nu(x_\nu) \, E_1(| z_\nu - x_\nu |) \, \mathrm{d}x_\nu \, . \tag{10.27}$$

Introducing this form for J_ν into formula (10.21), we obtain for S_ν the inhomogeneous integral equation

$$\begin{aligned} (\varkappa_\nu + \varkappa_\nu^0 + \sigma_\nu^0) \, S_\nu(z_\nu) &= (\varkappa_\nu + \varkappa_\nu^0) \, B_\nu(z_\nu) + \\ &+ \tfrac{1}{2} \sigma_\nu^0 \int_0^{\infty} S_\nu(x_\nu) \, E_1 (| z_\nu - x_\nu |) \, \mathrm{d}x_\nu \, . \end{aligned} \tag{10.28}$$

For faint lines, and for the outer parts of the wings of medium and strong lines, $\varkappa_\nu^0$ and σ_ν^0 are considerably less than $\varkappa_\nu$. Consequently,

the second term on the right-hand side of (10.28) represents a small correction. Hence we can introduce in this term (under the integral sign) the *approximate value* of S_ν, obtained by putting this second term on the right of (10.28) equal to zero. That is, we introduce in the integrand the following expression for S_ν:

$$S_\nu(z_\nu) = \frac{\varkappa_\nu + \varkappa_\nu^0}{\varkappa_\nu + \varkappa_\nu^0 + \sigma_\nu^0} B_\nu(z_\nu) . \tag{10.29}$$

Taking into account the fact that in our case $\varkappa_\nu^0 \ll \varkappa_\nu$ and $\sigma_\nu^0 \ll \varkappa_\nu$, we can assume that $S_\nu(z_\nu) \approx B_\nu(z_\nu)$. Using this result, we can rewrite (10.28) as

$$\begin{aligned}(\varkappa_\nu + \varkappa_\nu^0 + \sigma_\nu^0)\, S_\nu(z_\nu) &= (\varkappa_\nu + \varkappa_\nu^0)\, B_\nu(z_\nu) + \\ &+ \tfrac{1}{2}\,\sigma_\nu^0 \int_0^\infty B_\nu(x_\nu)\, E_1(\,|\,z_\nu - x_\nu\,|\,)\, \mathrm{d}x_\nu .\end{aligned} \tag{10.30}$$

We shall now calculate the flux $H_\nu(0)$ at the surface of the star, which is the quantity in which we are interested. From formulae (2.5) and (3.36), in which we must put $\tau_\nu = 0$ and introduce S_ν instead of B_ν, we find

$$\begin{aligned}\tfrac{1}{2}\, H_\nu(0) &= \int_0^{\frac{1}{2}\pi} I_\nu(0,0) \cos\theta \sin\theta\, \mathrm{d}\theta \\ &= \int_0^{\frac{1}{2}\pi} \cos\theta \sin\theta\, \mathrm{d}\theta \int_0^\infty S_\nu(z_\nu)\, e^{-z_\nu \sec\theta} \sec\theta\, \mathrm{d}z_\nu .\end{aligned} \tag{10.31}$$

Changing the order of integration and introducing a new variable w by means of the relation $\sec\theta = w$, we obtain, using (6.35),

$$\tfrac{1}{2}\, H_\nu(0) = \int_0^\infty S_\nu(z_\nu)\, E_2(z_\nu)\, \mathrm{d}z_\nu . \tag{10.32}$$

Substituting for $S_\nu(z_\nu)$ from (10.30) in (10.32), we obtain

$$\begin{aligned}\tfrac{1}{2}\, H_\nu(0) = \int_0^\infty & \frac{\varkappa_\nu + \varkappa_\nu^0}{\varkappa_\nu + \varkappa_\nu^0 + \sigma_\nu^0} B_\nu(z_\nu)\, E_2(z_\nu)\, \mathrm{d}z_\nu + \\ &+ \tfrac{1}{2} \int_0^\infty \frac{\sigma_\nu^0}{\varkappa_\nu + \varkappa_\nu^0 + \sigma_\nu^0} \left\{ \int_0^\infty B_\nu(x_\nu)\, E_1(\,|\,z_\nu - x_\nu\,|\,)\, \mathrm{d}x_\nu \right\} E_2(z_\nu)\, \mathrm{d}z_\nu .\end{aligned} \tag{10.33}$$

In (10.33) we carry out the following transformations. In the first integral we introduce instead of dz_ν the differentials $d\tau_\nu$ and $d\tau_\nu^0$:

$$\frac{\varkappa_\nu + \varkappa_\nu^0}{\varkappa_\nu + \varkappa_\nu^0 + \sigma_\nu^0}\,dz_\nu = \varkappa_\nu\,\varrho\,dh + \varkappa_\nu^0\,\varrho\,dh = d\tau_\nu + d\tau_\nu^0\,. \quad (10.34)$$

Next, we use the following differential relation, which follows from the definition (6.35):

$$dE_n(x)/dx = -E_{n-1}(x)\,. \quad (10.35)$$

From this expression, we can write

$$E_2(x+\Delta x) = E_2(x) - E_1(x)\,\Delta x\,. \quad (10.35a)$$

Introducing now the optical thickness $\tau_{\nu,\sigma}$ by means of the relation $d\tau_{\nu,\sigma} = \sigma_\nu^0\,\varrho\,dh$ and taking into account the fact that in our case $\tau_\nu^0 + \tau_{\nu,\sigma} \ll \tau_\nu$, we obtain from (10.35a)

$$E_2(z_\nu) = E_2(\tau_\nu + \tau_\nu^0 + \tau_{\nu,\sigma}) = E_2(\tau_\nu) - (\tau_\nu^0 + \tau_{\nu,\sigma})\,E_1(\tau_\nu) + \ldots\,. \quad (10.36)$$

This expression for $E_2(z_\nu)$ is substituted in the first integral in formula (10.33), and equation (10.34) is also used. Next, in the second integral in formula (10.33), we take into account the definition (10.23), and write τ_ν instead of z_ν, since $\tau_\nu^0 + \tau_{\nu,\sigma} \ll \tau_\nu$. As a result, we obtain

$$\begin{aligned}\tfrac{1}{2}\,H_\nu(0) &= \\ &= \int_0^\infty B_\nu(\tau_\nu)\,E_2(\tau_\nu)\,d\tau_\nu - \int_0^\infty B_\nu(\tau_\nu)\,E_1(\tau_\nu)\,(\tau_\nu^0 + \tau_{\nu,\sigma})\,d\tau_\nu + \qquad (10.37)\\ &+ \int_0^\infty B_\nu(\tau_\nu)\,E_2(\tau_\nu)\,d\tau_\nu^0 + \int_0^\infty \tfrac{1}{2}\,E_2(\tau_\nu)\,d\tau_{\nu,\sigma}\int_0^\infty B_\nu(x_\nu)\,E_1(|\tau_\nu - x_\nu|)\,dx_\nu\,.\end{aligned}$$

In transforming the first integral in formula (10.33) we have neglected, for reasons already given, the second term in the second parenthesis in the expression $(\tau_\nu^0 + \tau_{\nu,\sigma})\,(\varkappa_\nu + \varkappa_\nu^0)\,\varrho\,dh$, i. e. we have written, instead of this expression, $(\tau_\nu^0 + \tau_{\nu,\sigma})\,d\tau_\nu$.

We now perform our last transformation of the expression for $H_\nu(0)$. In calculating the second integral on the right-hand side of (10.37), we introduce the following function of τ_ν:

$$\Phi(\tau_\nu) = \int_{\tau_\nu}^\infty B_\nu(\tau_\nu)\,E_1(\tau_\nu)\,d\tau_\nu\,. \quad (10.38)$$

Integration by parts then gives

$$\int_0^\infty \tau_{\nu,\sigma} B_\nu(\tau_\nu) E_1(\tau_\nu)\,d\tau_\nu = -\left[\tau_{\nu,\sigma}\Phi(\tau_\nu)\right]_0^\infty + \int_0^\infty \Phi(\tau_\nu)\,d\tau_{\nu,\sigma} = \int_0^\infty \Phi(\tau_\nu)\,d\tau_{\nu,\sigma}\,. \tag{10.39}$$

We obtain a similar expression for the term in τ_ν^0. As a result of these transformations, the expression (10.37) takes the form

$$\begin{aligned}\tfrac{1}{2} H_\nu(0) = \int_0^\infty B_\nu(\tau_\nu) E_2(\tau_\nu)\,d\tau_\nu - \int_0^\infty [\Phi(\tau_\nu) - B_\nu(\tau_\nu) E_2(\tau_\nu)]\,d\tau_\nu^0 - \\ - \int_0^\infty \{\Phi(\tau_\nu) - \tfrac{1}{2} E_2(\tau_\nu) \int_0^\infty B_\nu(x_\nu) E_1(|\tau_\nu - x_\nu|)\,dx_\nu\}\,d\tau_{\nu,\sigma}.\end{aligned} \tag{10.40}$$

Putting $\varkappa_\nu^0$ and σ_ν^0 equal to zero, we obviously obtain the flux $H_\nu^0(0)$ for *the continuous radiation*, i. e. for the case where there is no selective scattering and absorption in the given frequency. We find from (10.40)

$$\tfrac{1}{2} H_\nu^0(0) = \int_0^\infty B_\nu(\tau_\nu) E_2(\tau_\nu)\,d\tau_\nu\,. \tag{10.41}$$

Now, writing

$$d\tau_\nu^0 = (\varkappa_\nu^0/\varkappa_\nu)\,d\tau_\nu \text{ and } d\tau_{\nu,\sigma} = (\sigma_\nu^0/\varkappa_\nu)\,d\tau_\nu\,, \tag{10.42}$$

we obtain for the line depth $R_\nu = 1 - r_\nu$ the following expression:

$$\begin{aligned}R_\nu = \frac{H_\nu^0(0) - H_\nu(0)}{H_\nu^0(0)} &= \int_0^\infty G_1(\tau_\nu)\,d\tau_\nu^0 + \int_0^\infty G_2(\tau_\nu)\,d\tau_{\nu,\sigma} \\ &= \int_0^\infty \frac{\varkappa_\nu^0}{\varkappa_\nu} G_1(\tau_\nu)\,d\tau_\nu + \int_0^\infty \frac{\sigma_\nu^0}{\varkappa_\nu} G_2(\tau_\nu)\,d\tau_\nu\,,\end{aligned} \tag{10.43}$$

where $G_1(\tau_\nu)$ and $G_2(\tau_\nu)$ are weight functions determined only by the dependence of B_ν on τ_ν:

$$G_1(\tau_\nu) = \frac{\Phi(\tau_\nu) - B_\nu(\tau_\nu) E_2(\tau_\nu)}{\int_0^\infty B_\nu(\tau_\nu) E_2(\tau_\nu)\,d\tau_\nu}\,, \tag{10.44}$$

$$G_2(\tau_\nu) = \frac{\Phi(\tau_\nu) - \frac{1}{2} E_2(\tau_\nu) \int_0^\infty B_\nu(x_\nu) E_1(|\tau_\nu - x_\nu|)\,dx_\nu}{\int_0^\infty B_\nu(\tau_\nu) E_2(\tau_\nu)\,d\tau_\nu}\,. \tag{10.45}$$

The dependence of $B_\nu(\tau_\nu)$ on τ_ν, and consequently the functions $G_1(\tau_\nu)$ and $G_2(\tau_\nu)$, are given by the theory of photospheres. For the Sun, this dependence can be found from the limb-darkening law (for various frequencies, see Chapter 7) established by observation. The ratios $\varkappa_\nu^0/\varkappa_\nu$ and $\sigma_\nu^0/\varkappa_\nu$ are known functions (see Chapter 11) of density, pressure, temperature, etc., while the distribution of these physical parameters is in turn known from the theory of photospheres. Thus, for given $G_1(\tau_\nu)$ and $G_2(\tau_\nu)$, the calculation of R_ν for a given frequency reduces to a simple calculation of the integrals in formula (10.43). This considerably facilitates the study of contours. However, it must be remembered that this method is applicable only to the wings of medium and strong lines and to faint lines (in general for cases where $R_\nu \leq 0{\cdot}15$).

The integral equation (10.28) for the function $S_\nu(z_\nu)$ can be used to calculate not only the flux $H_\nu(0)$ at the boundary, but also the emergent intensity $I_\nu(\theta,0)$, i.e. the depth $R_\nu(\theta)$ corresponding to an angle of emergence θ.

Besides the methods of solution of the equations of transfer (9.29), (9.30) and (9.32) which have been explained here, a number of other methods have been proposed by various authors. The majority of these methods involve simplifying physical or mathematical assumptions, and hence are of limited interest. They are most frequently concerned with assumptions about the linear expansion of the function B_ν and the constancy of the quantities η_ν and ε_ν. All these assumptions are, as we have said earlier, far removed from reality.

Recently, the iterational method of solving the equations of transfer, both for absorption lines and for frequencies in the continuous spectrum, has been widely used. A description of this method is given by, amongst others, L. H. Aller [2].

Chapter 11. The coefficients of selective absorption

1. **The relation between the absorption coefficients and the transition coefficients.** In this chapter we shall consider the coefficients of selective absorption, i.e. the absorption coefficients inside spectral lines. This subject amounts to an examination of the factors which bring about the *finite width* of absorption lines in stellar (and solar) spectra. The coefficients of selective absorption are functions which give the distribution of the total number of absorption processes $n_i B_{ik} \varrho_{\nu_{ik}}$ in a finite frequency interval.

With a view to future applications, we shall now derive an integral formula for the absorption coefficient, which is of wide use in astrophysics. Let n_i and n_k be the number of atoms in the ith and kth excitation states respectively in 1 cm³. Also, let the atomic selective absorption

coefficient corresponding to the transition $i \to k$ be s_ν. Let the radiation incident on the atoms considered have a density ϱ_ν. We shall assume that, within the spectral region (generally very narrow) where s_ν is fairly large, the density of radiation varies only slightly with frequency. We recall that this condition must be satisfied for the transition coefficients A_{ik}, B_{ik}, B_{ki} to be applicable.

Let us write down expressions for the number of transitions $i \to k$ which take place in 1 cm^3 in 1 second. This number is determined, in terms of the transition coefficients, by formula (5.38):

$$n_{i\to k} = n_i \, \varrho_{\nu_{ik}} \, B_{ik} \, . \tag{11.1}$$

We now write an expression for $n_{i\to k}$ using the coefficient s_ν. The number of absorptions $(n_{i\to k})_\nu \, \mathrm{d}\nu$ inside the frequency interval $\mathrm{d}\nu$ can be written, from formula (5.47), as

$$(n_{i\to k})_\nu \, \mathrm{d}\nu = n_i \, s_\nu \, \frac{c \, \mathrm{d}\nu}{h \, \nu} \, \varrho_\nu \, . \tag{11.2}$$

In order to obtain $n_{i\to k}$, this equation must be integrated over the frequency. As a result we find

$$n_{i\to k} = \int (n_{i\to k})_\nu \, \mathrm{d}\nu = n_i \, \frac{c}{h} \int \varrho_\nu \, \frac{s_\nu}{\nu} \, \mathrm{d}\nu \, . \tag{11.3}$$

Next, because of the narrowness of the region where s_ν is fairly large, we can take ϱ_ν and ν outside the integral in (11.3). This gives

$$n_{i\to k} = n_i \, \frac{c}{h \, \nu_{ik}} \, \varrho_{\,\nu_{ik}} \int s_\nu \, \mathrm{d}\nu \, . \tag{11.4}$$

Identifying (11.1) and (11.4), we obtain

$$\int s_\nu \, \mathrm{d}\nu = \frac{h \, \nu_{ik}}{c} \, B_{ik} \, . \tag{11.5}$$

According to this formula, the integral of the absorption coefficient is a constant quantity, and is consequently independent of the mechanism which broadens the line in question.

We must now acquaint ourselves with the chief processes which determine the form of s_ν. Here we shall use classical ideas as well those of quantum theory.

2. **Radiation damping.** Let us consider the process of the damping of a classical atomic oscillator *by radiation.* The energy of a radiating oscillator, according to classical electrodynamics, diminishes according to the law

$$E = E_0 \, e^{-\gamma_0 t} \, , \tag{11.6}$$

where

$$\gamma_0 = 8\pi^2 e^2 \nu_0^2/3 m_e c^3 \tag{11.7}$$

is called the **damping decrement.** The waves emitted by such a damped oscillator do not produce an infinitely narrow spectral line, but one of finite width. Strictly monochromatic radiation (i. e. an absolutely sharp emission line) is obtained only when we deal with *an infinite wave train of constant amplitude.*

The actual frequency distribution of the radiation is obtained by analysing the oscillation in question into purely harmonic components (using a Fourier integral). Correspondingly, the absorption process is also non-monochromatic. The absorbed energy is distributed in a definite manner within a finite frequency interval. The classical theory of dispersion gives the following expression for the selective absorption coefficient:

$$s_\nu = \frac{e^2}{m_e c} \frac{\gamma_0}{4\pi} \frac{1}{(\nu - \nu_0)^2 + (\gamma_0/4\pi)^2}, \tag{11.8}$$

where ν_0 is the **natural frequency** of the oscillator vibrations, and γ_0 is the damping constant determined by formula (11.7).

The absorption coefficient which we have written is referred to one oscillator. Further, the formula (11.8) is valid only for a fairly narrow frequency interval inside an absorption line, where ν is close to ν_0. (We notice that the same dependence on frequency is obtained for the process of emission by the oscillator.)

The quantity γ_0 is called the **natural line width.** This name relates to the fact that, in the *scale of circular frequencies* $\omega = 2\pi\nu$, the absorption coefficient, as follows from (11.8), diminishes by half at a distance $\omega - \omega_0 = \frac{1}{2}\gamma_0$. Doubling this quantity to take account of the two halves of the contour of s_ν in both directions from the frequency ν_0, we obtain γ_0. Hence γ_0 is sometimes called also the **total half-width of the absorption coefficient.** If the value of γ_0 is expressed in the wavelength scale by means of (11.7) and (4.38), it is independent of λ and equal to $1{\cdot}18 \times 10^{-4}$ Å.

The quantity τ_0 which is the reciprocal of the damping constant γ_0 is called the **damping time** of the oscillator. The reason for this definition is clear from formula (11.6). For the sodium lines D_1 and D_2, whose mean wavelength $\lambda = 5893$ Å, the damping time calculated by means of (11.7) is $1{\cdot}58 \times 10^{-8}$ sec.

At fairly large pressures, the damping of the oscillator *by collisions* begins to play an important part, as well as radiation damping. In each collision, the vibration of the oscillator is interrupted for a moment, and almost at once a new vibration is excited with *arbitrary phase and amplitude.* In other words, the collisions cut short the sequence of emitted waves and thereby also increase the non-monochromatic character of the spectral line.

The appropriate analysis [24, § 87] shows that collision damping (as yet without taking account of radiation damping) leads to an absorption coefficient of exactly the same form as (11.8). However, in this case γ_0 must be replaced by a new damping constant, the collision damping constant γ_c. This constant is related to the mean time τ_c between two collisions by

$$\gamma_c = 2/\tau_c \tag{11.9}$$

(whereas for radiation damping we have the relation $\gamma_0 = 1/\tau_0$).

If, finally, we take account of simultaneous radiation and collision damping, then the form of the absorption coefficient still remains the same [formula (11.8)], but it is necessary to introduce into this formula the effective damping constant γ, defined by the equation

$$\gamma = \gamma_0 + \gamma_c \,. \tag{11.10}$$

Collisions in many cases play a very great part in the formation of absorption lines in stellar spectra.

In all the cases mentioned the absorption coefficient s_ν, given by formulae similar to (11.8), was referred to one oscillator. To pass from the number of oscillators $\bar{n}$ to the number of atoms n in the state from which the transitions begin which form the given absorption line, we introduce what is called the **oscillator strength** f:

$$\bar{n} = f\, n \,. \tag{11.11}$$

Thus *the oscillator strength f gives the number of classical oscillators which reproduce the absorbing action of one atom in the given line.* It turns out that in this case the absorption coefficient for one atom is expressed by a formula of the same type as (11.8), with the difference that $\gamma_0 + \gamma_c$ must be written in place of γ_0, and the whole expression must be multiplied by the oscillator strength corresponding to the given line:

$$s_\nu = \frac{e^2}{m_e c}\,\frac{\gamma}{4\pi}\,\frac{f}{(\nu - \nu_0)^2 + (\gamma/4\pi)^2} \,. \tag{11.12}$$

If we integrate both sides of (11.12) over frequency from 0 to $+\infty$, we find

$$\int_0^\infty s_\nu \,\mathrm{d}\nu = (\pi e^2/m_e c)\, f \,. \tag{11.13}$$

Comparing the expressions (11.5) and (11.13), we obtain

$$f_{ik} = (m_e h \nu_{ik}/\pi e^2)\, B_{ik} \,, \tag{11.14}$$

or, using (5.40) and (11.7),

$$f_{ik} = \frac{g_k}{g_i} \frac{m_e c^3}{8\pi^2 e^2 \nu_{ik}^2} A_{ki} = \frac{g_k}{g_i} \frac{A_{ki}}{3\gamma_0} . \tag{11.15}$$

In order to emphasise that we are speaking of the quantum-theory transitions $i \rightarrow k$ of the atom, we have added the symbol ik to the oscillator strength f.

Let us now consider the question of the broadening of absorption lines according to the ideas of contemporary quantum mechanics. Let there be n_k excited atoms in 1 cm^3. Neglecting stimulated transitions (both downwards and upwards from the level k) we can write the obvious expression for the decrease of n_k with time

$$\mathrm{d}n_k/\mathrm{d}t = -n_k \sum_i A_{ki} . \tag{11.16}$$

Integrating this expression, we find

$$n_k = (n_k)_0 \, e^{-\gamma_k t} , \tag{11.17}$$

where $(n_k)_0$ is the value of n_k for $t = 0$, while

$$\gamma_k = \sum_i A_{ki} = 1/\tau_k . \tag{11.18}$$

If we multiply both sides of (11.17) by $\sum_i A_{ki} h \nu_{ik}$, the law (11.17) which gives the decrease of the total excitation energy is wholly similar to the classical law (11.6). Hence we can call γ_k here also the **damping constant**, and τ_k the **damping time**.

The quantity τ_k is also the **mean life** of the atom in the state k (or simply the "mean life of the state k"). In fact, the probability that the mean duration of the atom's existence in the excited state k lies between t and $t + \mathrm{d}t$ is proportional, according to (11.17), to the differential $e^{-\gamma_k t} \gamma_k \, \mathrm{d}t$. Consequently, the (weighted) mean life $\bar{t}_k$ of the atom in the state k is

$$\bar{t}_k = \int_0^\infty t e^{-\gamma_k t} \gamma_k \, \mathrm{d}t \Big/ \int_0^\infty e^{-\gamma_k t} \gamma_k \, \mathrm{d}t = 1/\gamma_k = \tau_k , \tag{11.19}$$

which is the result required.

The values of τ_k and of γ_k can be found from formula (11.18) if the A_{ki} are known for all downward transitions. If the atoms are in a fairly strong radiation field (high temperature of the radiation), it is necessary to take account not only of the spontaneous transitions downwards

from the level k, but also of the transitions, both upwards and downwards (stimulated emission), due to the action of the radiation. As a result, we can write for τ_k and γ_k, instead of (11.18),

$$\gamma_k = 1/\tau_k = \sum_i A_{ki} + \sum_i \varrho_{\nu_{ik}} B_{ki} + \sum_\sigma \varrho_{\nu_{\sigma k}} B_{k\sigma} + Z^k_{pi}\,. \tag{11.20}$$

The first sum on the right-hand side of (11.20) takes account of *spontaneous* transitions downwards from the level k, the second takes account of *stimulated* transitions, also downwards, the third gives all possible (*discrete*) transitions upwards from the level k, and finally the last term on the right-hand side of (11.20) gives the *photo-ionisation* transitions from the level k. The latter must be calculated in accordance with the expression (8.6), which of course must be first divided by $n_{r,\,k}$.

The finiteness of the mean life of the atom in the excited state leads, in accordance with what is called the **uncertainty principle**, to a broadening of the energy level in question. Here, as is seen from the well-known formula

$$\Delta E\,\Delta t \approx h/2\,\pi\,, \tag{11.21}$$

the broadening of any level is the greater, the shorter the mean life of the atom for this level.

If the excitation of the atom to the state k takes place under the action of light with a constant intensity in the given frequency interval, or under the action of collisions with particles having all possible velocities, then, for a finite value of the mean life of the atom, the probability that an atom in the state k has an energy between E and $E + \mathrm{d}E$ is

$$W(E)\,\mathrm{d}E = \frac{\gamma_k}{h}\,\frac{\mathrm{d}E}{(2\,\pi/h)^2\,(E - E_k)^2 + (\gamma_k/2)^2}\,. \tag{11.22}$$

The normalisation of (11.22) is such that $\int W(E)\,\mathrm{d}E = 1$.

If now we are concerned with transitions between the levels i and k, where the distribution (11.22) holds for each level, then it can be shown [168, p. 282] that the relative probability of the emission or absorption of a quantum in the frequency range from ν to $\nu + \mathrm{d}\nu$ is

$$P_\nu\,\mathrm{d}\nu = \frac{(\gamma_k + \gamma_i)\,\mathrm{d}\nu}{4\,\pi^2\,(\nu - \nu_{ik})^2 + (\gamma_k + \gamma_i)^2/4}\,. \tag{11.23}$$

Consequently, we have the same distribution as in the case of (11.12), but with $\gamma_0 = \gamma_i + \gamma_k$.

Finally, taking account of the normalising equation (11.13) for the coefficient s_ν, which is proportional to P_ν, we arrive at the result that the same classical expression (11.12) can be used for s_ν, but the value

of γ_0 must be taken as the sum of the constants γ_i and γ_k. Hence, from the standpoint of quantum theory, we can write for s_ν

$$s_\nu = \frac{e^2}{m_e c}\,\delta_{ik}\,\frac{f}{(\nu - \nu_{ik})^2 + \delta_{ik}^2} = \frac{c^2}{8\,\pi^2\,\nu_{ik}^2}\,\frac{g_k}{g_i}\,\frac{A_{ki}\,\delta_{ik}}{(\nu - \nu_{ik})^2 + \delta_{ik}^2}\,, \tag{11.24}$$

where

$$\delta_{ik} = \frac{\Gamma_{ik}}{4\,\pi} = \frac{1}{4\,\pi}\,(\gamma_i + \gamma_k + \gamma_c) = \frac{1}{4\,\pi}\left(\frac{1}{\tau_i} + \frac{1}{\tau_k} + \frac{2}{\tau_c}\right). \tag{11.25}$$

Thus the existence of damping in the lower state is also taken into account in the quantum formula.

If the lower level i is the ground level of the atom, its mean life is large, and consequently the constant γ_i is small; hence the ground level may be considered sharp (if, of course, the radiation field is not too intense in the frequencies of the transition from the lower to the higher state). Thus the first term in parentheses in (11.25) disappears for lines which start from the ground level. Fig. 24 shows a schematic representation of the ground level and two excited levels k and l. The energy broadening of the levels k and l is shown in the figure by the sets of parallel lines.

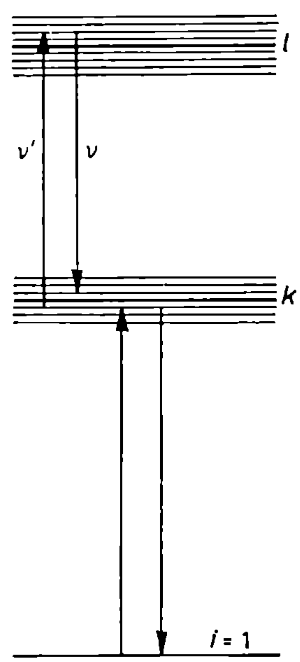

Fig. 24

We must now notice the following important circumstance. It follows from Chapter 9 that, besides the processes of selective absorption, those of re-emission of absorbed energy also play a very important part in the theory of stellar absorption lines. It was supposed in Chapter 9, and therefore in Chapter 10, that we are concerned with processes of coherent re-emission, where the absorbed quantum corresponds to a re-emitted quantum of exactly the same frequency. It is quite evident that this condition cannot be satisfied for transitions between excited levels. In fact, it follows from Fig. 24 that the absorption of a quantum of frequency ν' can lead to the re-emission of a quantum of frequency ν, which is different from ν', though close to it. In this case the equations of transfer (9.6), (9.29), (9.30) and (9.32) and the fundamental expression for the coefficient of emission (scattering) cease to be valid. It follows from Fig. 24 that the coherence of emission necessarily holds only for a resonance line, i. e. for a line whose lower level is the ground state, and whose upper level is that closest to the ground state. In this case the frequencies of the absorbed and emitted quanta must be equal. It is true, as we shall see in Chapter 14, that

even here there are effects which destroy the coherence. However, the deviations from coherence are least for resonance lines. The whole of the present-day theory of absorption lines has been developed for coherent re-emission. Hence we shall, for the time being, be concerned with this case, and consequently we shall be discussing mainly resonance lines, or at any rate lines starting from the ground level. In this last case the coherence may be destroyed because, for instance, an electron which has been raised from the ground level to some excited level l falls back from there not to the ground level, but to some lower excited level k (see Fig. 24), and the frequencies of the absorbed and emitted quanta will simply correspond to different lines of the atom concerned [168, § 71].

3. **Thermal and turbulent motions of atoms.** Let us consider another process which leads to the broadening of lines, namely the Doppler effect brought about by the motion of the absorbing and emitting atoms. Here there are two possibilities: (1) the thermal motion of the atoms: since a large number of atoms always takes part in the absorption of radiation, and their thermal velocities are very different in both magnitude and direction, this would inevitably lead, even if damping by radiation and collisions were absent, to a broadening of absorption lines; and (2) as we shall see in the next chapter, the material in stellar atmospheres (a gas) is characterised by the presence of irregular large-scale motions in every direction; apparently, gaseous masses (separate condensations) of the most various dimensions take part in these motions. These motions have acquired in astrophysics the name of "turbulent" motions, although the nature of this "turbulence" may in some cases differ from that of the ordinary turbulence considered in hydrodynamics. It is quite evident that these motions also, like the thermal motions, lead to a broadening of absorption lines.

Let us consider at first only the thermal Doppler effect, assuming that the effects of radiation damping and collision damping are absent, i. e. in formula (11.24) $\delta_{ik} = 0$. Let the component of the atom's velocity in the line of sight be v. Then, in the presence of a Maxwellian velocity distribution, the proportion of atoms whose radial velocities lie between v and $v + dv$ is

$$\frac{dn}{n} = \frac{1}{\sqrt{\pi}} e^{-(v/v_0)^2} dv/v_0 \tag{11.26}$$

(the integral of the right-hand side of (11.26) over all v from $-\infty$ to $+\infty$ is equal to unity), where $\frac{1}{2} v_0^2$ is the mean value of v^2:

$$\overline{v^2} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} v^2 e^{-(v/v_0)^2} dv/v_0 = \tfrac{1}{2} v_0^2 . \tag{11.27}$$

The quantity v_0 is related to the temperature of the gas by

$$v_0^2 = 2\,RT/\mu\,, \tag{11.28}$$

where μ is the atomic weight and R the gas constant.

The law of distribution of "turbulent" velocities within stellar atmospheres is not yet completely known. If, postulating chaotic irregular "turbulent" motions, we assume that the distribution law (11.26) with $v_0 = v_t$ holds here also, then, according to the rule of superposition of dispersions, the resultant velocity distribution of atoms will be of the same form as (11.26), but with

$$v_0^2 = 2\,RT/\mu + v_t^2\,. \tag{11.29}$$

According to Doppler's principle, which was first experimentally tested by Academician A. A. BELOPOL'SKIĬ, the frequency or wavelength displacement from the normal value is determined by the law

$$\frac{\Delta\nu}{\nu} = \frac{\nu - \nu_0}{\nu} = \frac{\Delta\lambda}{\lambda} = \frac{\lambda - \lambda_0}{\lambda} = \frac{v}{c}\,, \tag{11.30}$$

where the absolute values of the differences $\lambda - \lambda_0$ and $\nu - \nu_0$ are here understood. Putting

$$\Delta\nu_D/\nu_0 = \Delta\lambda_D/\lambda_0 = v_0/c\,, \tag{11.31}$$

we obtain the following expression for the relative frequency distribution, inside the line, of absorbing and emitting atoms:

$$\left.\begin{aligned} \frac{\mathrm{d}n}{n} &= \frac{1}{\sqrt{\pi}}\, e^{-(\Delta\nu/\Delta\nu_D)^2}\,\mathrm{d}\nu/\Delta\nu_D\,, \\ \text{or}\qquad & \\ \frac{\mathrm{d}n}{n} &= \frac{1}{\sqrt{\pi}}\, e^{-(\Delta\lambda/\Delta\lambda_D)^2}\,\mathrm{d}\lambda/\Delta\lambda_D\,, \end{aligned}\right\} \tag{11.32}$$

where we have taken into account that $\mathrm{d}(\Delta\nu) = \mathrm{d}\nu$ and $\mathrm{d}(\Delta\lambda) = \mathrm{d}\lambda$. For purely thermal motion

$$\left.\begin{aligned} \Delta\nu_D &= \frac{\nu}{c}\sqrt{\left(\frac{2\,RT}{\mu}\right)} = \frac{\nu}{c}\sqrt{\left(\frac{2\,kT}{m}\right)} \\ \text{and}\qquad & \\ \Delta\lambda_D &= \frac{\lambda}{c}\sqrt{\left(\frac{2\,RT}{\mu}\right)} = \frac{\lambda}{c}\sqrt{\left(\frac{2\,kT}{m}\right)}, \end{aligned}\right\} \tag{11.33}$$

where k is Boltzmann's constant and m is the mass of the atom considered.

Taking into consideration the proportionality between dn/n and the absorption coefficient s_ν, we can normalise the latter by means of formulae (11.5) and (11.14); here the integration can be taken from $\nu - \nu_0 = -\infty$ to $\nu - \nu_0 = +\infty$, since even for relatively small $(\nu - \nu_0)/\Delta\nu_D$ the exponential factor tends rapidly to zero, and the change of limits of integration does not involve a noticeable error. As a result of the integration we obtain

$$s_\nu = \frac{\sqrt{\pi}\, e^2}{m_e c} \frac{f_{ik}}{\Delta\nu_D} e^{-(\nu-\nu_0)^2/(\Delta\nu_D)^2} \tag{11.34}$$

A similar expression is obtained for s_λ.

The quantities $\Delta\nu_D$ and $\Delta\lambda_D$ are called **Doppler widths.** For $\nu - \nu_0 = \Delta\nu_D$ or $\lambda - \lambda_0 = \Delta\lambda_D$, the absorption coefficient diminishes by a factor e. It can easily be calculated from formula (11.33) that, for a temperature equal to the effective temperature of the Sun (5710°), the Doppler width $\Delta\lambda_D$ for the H_α line of hydrogen is 0·212 Å, for the D_1 line of neutral sodium ($\lambda = 5895{\cdot}9$ Å) it is 0·0399 Å, and for the K line of ionised calcium ($\lambda = 3933{\cdot}7$ Å) it is 0·0202 Å. At 10,000° the values for the same lines are respectively 0·281, 0·0529 and 0·0267. Thus these quantities are considerably larger than the natural widths of the lines, which are of the order of 10^{-4} Å.

4. **The simultaneous action of radiation damping and the motion of atoms.** A direct application of (11.34) to determine the absorption in some gaseous layer is possible only if the number of absorbing atoms is not too large. When the number of atoms increases, the broadening by damping begins to play a part. Thus, in the general case, it is necessary to take both effects into account at once (the damping effect and the Doppler effect).

The absorption coefficient in the general case can be obtained as follows. Let us assume that an atom having an absorption coefficient determined by formula (11.24) is moving with velocity v relative to the observer. In this case the central frequency of the absorption process ν_{ik} is displaced by a quantity $\Delta\nu$ determined by formula (11.30). For the observer, the absorption coefficient in the frequency ν referred to one such atom is

$$\frac{e^2}{m_e c} \delta_{ik} \frac{f_{ik}}{(\nu - \nu_{ik} - \Delta\nu)^2 + \delta_{ik}^2} . \tag{11.35}$$

On the other hand, the relative proportion of such atoms is, according to (11.32),

$$\frac{1}{\sqrt{\pi}} e^{-(\Delta\nu/\Delta\nu_D)^2} \, d\Delta\nu/\Delta\nu_D , \tag{11.36}$$

since $d\nu = d(\nu - \nu_{ik}) = d\Delta\nu$.

Consequently, the absorbing effect produced by the atoms in question is determined by the product of the expressions (11.35) and (11.36). In order to obtain the resultant absorption coefficient, we must take into account the fact that the absorption in the frequency ν is due not only to these atoms which move with radial velocity v, but to all the atoms, having all possible velocities (relative to the observer) from $v = -\infty$ to $v = +\infty$. Adding the absorbing effect of all these atoms, we obtain

$$s_\nu = \frac{e^2 \delta_{ik} f_{ik}}{\sqrt{\pi}\, m_e c} \int_{-\infty}^{\infty} \frac{e^{-(\Delta\nu/\Delta\nu_D)^2} \, d\Delta\nu/\Delta\nu_D}{(\nu - \nu_{ik} - \Delta\nu)^2 + \delta_{ik}^2} \,. \tag{11.37}$$

Here, of course, we have used the fact that the integral of (11.36) over all $\Delta\nu$ is unity. Next, we put

$$y = \Delta\nu/\Delta\nu_D, \; p = (\nu - \nu_{ik})/\Delta\nu_D, \; a = \delta_{ik}/\Delta\nu_D \,. \tag{11.38}$$

Then (11.37) takes the form

$$\frac{s_\nu}{s_{\nu_0}} = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-y^2} \, dy}{a^2 + (p - y)^2} = \phi(a,p) \,, \tag{11.39}$$

where

$$s_{\nu_0} = \frac{\sqrt{\pi}\, e^2}{m_e c} \frac{f_{ik}}{\Delta\nu_D} \,, \tag{11.40}$$

and, in accordance with (11.34), s_{ν_0} is equal to the central value (for $\nu = \nu_0$) of the absorption coefficient due to the Doppler effect alone.

It follows from formula (11.39) that the problem of determining s_ν as a function of ν reduces to the calculation of the integral on the right-hand side of this formula, i. e. the function $\phi(a,p)$. We shall not treat this integral in its general form. For values of a considerably less than unity ($a \ll 1$), it can be evaluated with sufficient accuracy by means of various expansions of the function $\phi(a,p)$ in series [168, p. 263]. We shall omit this analysis, and give only the final results. These are important because the case $a \ll 1$ holds for the great majority of the absorption lines formed in stellar atmospheres.

The most important of these results are as follows. For $a \ll 1$, the absorption coefficient given by formula (11.39) may be approximately represented in different parts of the line by two different expressions: the first is determined only by the Doppler effect [in accordance with formula (11.34)], and the second only by the damping [formula (11.24)].

The expression (11.34) plays the chief part in the central frequencies close to $\nu = \nu_{ik}$; the second expression (determined by the damping) does so for frequencies more distant from the centre, where the Doppler broadening does not play any part [this follows from the exponential character of the law (11.34)]. In the intermediate part which joins these two regions, the absorption coefficient cannot be represented either by formula (11.24) or by formula (11.34)*.

For the case $a \ll 1$, the value of s_{ν_0} is in fact the central value of s_ν, calculated from (11.40). All these results are illustrated in Fig. 25, where the graph of $-\log_{10}(s_\nu/s_{\nu_0})$ is shown (left-hand ordinate scale) for the transitions which produce the sodium line D_1. The values of the absorption coefficient calculated for damping only and for Doppler broadening only are shown dotted. Here T is taken as 5700°. The values of $|\lambda - \lambda_0|$ in Å are placed on the axis of abscissae at the top, and the values of p from formula (11.38) at the bottom. The values of $-\log_{10}(s_\nu/f)$, i.e. the absorption coefficient referred to one oscillator, are given on the axis of ordinates at the right. Calculations show† that at this temperature $a = 3{\cdot}0 \times 10^{-3}$.

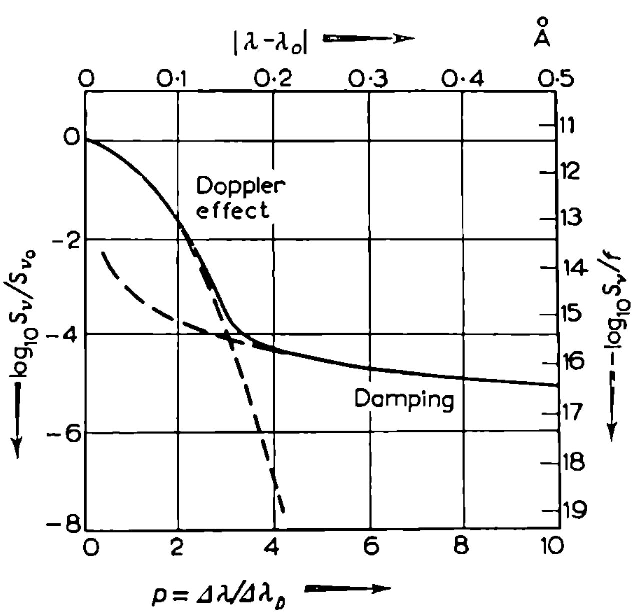

FIG. 25

We may make the following remarks concerning Fig. 25, which is valid generally for the case $a \ll 1$. Firstly, it follows from this figure that the Doppler core almost entirely determines the value of the integral on the left-hand side of the fundamental expression (11.5). Secondly, the intermediate joining part of the curve is narrow for the case $a \ll 1$, so that we can use either formula (11.24) or formula (11.34) in practically the whole frequency interval.

In cases where a does not satisfy the inequality $a \ll 1$, it is necessary to use the tables of the function $\phi(a, p)$ which are given at the end of the book (see Appendix I).

When a increases, the representation of s_ν in two parts, the Doppler part and the damping part, becomes less and less accurate. This is expressed in the fact that the intermediate joining part of the s_ν curve

* Tables for the calculation of the function $\phi(a, p)$ are given in Appendix I at the end of this book.

† Only radiation damping is taken into account.

becomes wider and wider as a increases, and the right-hand side of formula (11.39) for $p = 0$, i. e. for the centre of the line, ceases to be equal to unity (though it does not differ very much from unity). Thus, it follows from the tables mentioned that, for instance, for $a = 0{\cdot}2$ the ratio s_ν/s_{ν_0} for $\lambda \approx \lambda_{ik}$ is 0·809, so that s_{ν_0} ceases to be the central value of s_ν.

However, in the majority of cases encountered in practice, a is very small, and hence we can construct s_ν in these cases in the manner shown above, i. e. from two expressions. Moreover, for practical calculations s_{ν_0} may be taken as the central value of s_ν. In general, the validity of this method must be tested in doubtful cases by means of the above-mentioned tables.

5. **Pressure effects.** To conclude the present chapter, we shall discuss the question of line broadening as a result of the interaction of the absorbing atoms with the surrounding particles. This form of line broadening has the general name of the **pressure effect,** since it is the greater (for a given temperature), the greater the pressure of the gas. As we shall see in the two following chapters, the main part in determining the value of δ_{ik} [see (11.25)] is very often played not by radiation damping (i. e. the sum $\gamma_i + \gamma_k$), but by collision damping, due to the collisions of the absorbing atoms with the particles surrounding them ($\gamma_c \gg \gamma_i + \gamma_k$). Moreover, for hydrogen lines and some helium lines, formula (11.24) is in general inapplicable by reason of the pressure effect.

The pressure effect includes the following two processes:

(1) Collisions in which the absorbing atom gives up its excitation energy to the surrounding particles without subsequent re-emission (collisions of the second kind).

(2) Collisions between absorbing atoms and neighbouring particles which lead to a broadening of the absorption line without loss of the excitation energy of the atoms.

In the first case, the collisions lead to a transformation of excitation energy into heat; thus *true selective absorption* takes place, which we considered at the end of Chapter 9. However, if the chief processes which determine the energy balance in absorption lines were processes of true absorption, then the absorption lines would disappear at the Sun's limb (see Chapter 9), and this is contrary to observation. Another form of loss of excitation energy is that by *collisions of excited atoms with molecules*. Here the excitation energy of the atom is used to excite the molecule which collides with it. The efficiency of processes of this kind is due to the great complexity of the term system of molecules. When a molecule collides with an excited atom whose excitation energy is ε_i,

there is a high probability that there is a transition in the molecule with the same value of ε_i. However, this case also cannot be important, since there are many more neutral hydrogen atoms, which produce line broadening (see below), than molecules in the atmospheres of cool stars. Hence the part played by hydrogen atoms is greater than that played by molecules. Having regard to these remarks, we shall in future consider only the second of the two cases stated, which causes line broadening without loss of the excitation energy of the atoms.

A physical analysis of the corresponding processes shows that the approach to the solution of this problem can be made in two ways. In one, processes are considered which we have already discussed, called *collision damping of atoms*. In the other, the direct influence of neighbouring particles on the energy level of the absorbing atom is considered. This influence is determined by the aggregate effect of many particles at various distances from the atom, and consequently it can be computed by statistical methods. The principal effect here considered is broadening by the intermolecular Stark effect (see below). The relation between the two aspects of the same pressure effect has been investigated only very recently. It has been shown that collision damping plays the principal part in the more central parts of absorption lines, while the "statistical" part of the effect operates in the outer parts. For given physical conditions, there is a certain critical distance $\Delta\lambda_g$ inside each line, measured from the centre of the line. Within the region $\pm\Delta\lambda_g$ the broadening is due to collision damping, but outside it the broadening is determined by statistical theory.

On the other hand, in various particular cases the conditions of interaction between atoms and other particles (or even between atoms of the same kind) may be totally different. For example, an interaction may take place between neutral atoms, or between neutral atoms and ions or electrons, etc. Hence the value of $\Delta\lambda_g$ may be quite different in different cases. In some cases $\Delta\lambda_g$ is relatively so small that the broadening can be described by statistical theory within almost the whole spectral line. In other cases $\Delta\lambda_g$ is so large that collision damping plays the principal part within the whole observed line; in this latter case statistical broadening begins to play an important part only for very distant parts of the wings, which cannot be recorded by observation, since $r_\nu \approx 1$ there.

Thus, in order to ascertain what type of interaction must be considered the main one, we have to determine the physical conditions exsiting in the given region of the stellar atmosphere, and also between which particles the interaction takes place.

We shall briefly discuss both forms of the pressure effect.

6. Line broadening by collisions. We have already, in deriving the constant γ_c, investigated the physical side of the question of collision broadening of lines. At every collision the oscillatory motion which is performed by the harmonic oscillator is momentarily curtailed, and begins again after a very short interval of time, though with a different phase. Thus, in constructing the theory, we must suppose that the wave train with frequency ν_0 undergoes from time to time a phase jump, the phase change at these jumps being arbitrary each time.

The starting point of the theory is the nature of the phase change when a perturbing particle passes near an absorbing (or emitting) atom. If $\Delta\nu$ is the change of frequency in the radiation of the atom at that moment, due to the passage of the perturbing particle past the atom, then the total phase change $\Delta\phi$ in the whole (very short) time of their mutual approach is

$$\Delta\phi = 2\,\pi \int \Delta\nu \,\mathrm{d}t\,, \tag{11.41}$$

where the integration is extended over the whole time of approach. The calculation of this integral is one of the main problems of the present theory.

A study of all physically possible cases shows that the value of $\Delta\nu$ as a function of the distance r between the absorbing (or emitting) atom and the perturbing particle can be expressed in the form

$$\Delta\nu = C/r^k\,, \tag{11.42}$$

where k is some integer and C a constant. The constant C can be determined from quantum-mechanical calculations or from laboratory measurements. In the case where the broadening particles are ions or electrons (with the exception of hydrogen lines and some helium lines), the change of frequency of the absorbing (or emitting) atom due to the ion or electron moving past it can be considered as a quadratic Stark effect (the splitting of the line is proportional to the square of the field strength $F = e/r^2$). Consequently, we can put

$$\Delta\nu = C/r^4\,, \tag{11.43}$$

so that in this case $k = 4$. Measuring the field strength in kilovolts (F in kilovolts $= F/0{\cdot}3$ in CGS electrostatic units) and the displacement $\Delta\lambda$ in Å, we find

$$\frac{c\,\Delta\lambda}{\lambda^2\cdot 10^{-8}} = \frac{C}{e^2}\left(\frac{e}{r^2}\right)^2 = \frac{C}{e^2}\left(\frac{F}{0{\cdot}3}\right)^2. \tag{11.44}$$

Thus, knowing the value of the splitting $\Delta\lambda$ and the corresponding field strength F from laboratory measurements, we can determine the constant C from (11.44).

Let us now consider the principal cases of collision damping which cause line broadening in stellar atmospheres. These cases correspond to $k = 3$, $k = 4$ and $k = 6$. We shall consider them in turn.

$k = 3$. In this case the perturbing particles are atoms of the same element as the one whose line broadening we are studying ("self-pressure broadening"). Here the following expression is obtained for γ_c:

$$\gamma_c = 4\,\pi^3\,C\,n\,, \tag{11.45}$$

where n is the number of perturbing particles of the given kind in 1 cm^3 [in this case, the absorbing (or emitting) atoms themselves]. The value of C is here *approximately*

$$C = \frac{e^2}{16\,\pi^2\,m_e\,\nu_0}\,f_{ik}\,, \tag{11.46}$$

where f_{ik} is the oscillator strength for the given transition and ν_0 is, as usual, the frequency of the line centre.

$k = 4$. We have already considered this case above. The perturbing particles are ions and electrons (quadratic Stark effect). For γ_c we obtain in this case

$$\gamma_c = 38{\cdot}8\,C^{\frac{2}{3}}\,\bar{v}^{\frac{1}{3}}\,n\,. \tag{11.47}$$

In this formula n is again the number of perturbing particles (in this case, electrons or ions) in 1 cm^3, and $\bar{v}$ is the mean relative velocity of the perturbed and perturbing particles. If the atomic weight of the former is μ_1, and of the latter μ_2, then by the kinetic theory of gases

$$\bar{v} = \sqrt{\left[\frac{8}{\pi}\,RT\left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right)\right]}\,, \tag{11.48}$$

werhe R is the gas constant and T the kinetic temperature of the medium. The values of C, as we have said, are either determined in the laboratory or calculated from the laws of quantum mechanics.

$k = 6$. In this case the line broadening is caused by the collisions of the perturbed atoms with those of neutral hydrogen. The forces of interaction for $k = 6$ are of the van der Waals type. Here γ_c has the form

$$\gamma_c = 17{\cdot}0\,C^{2/5}\,\bar{v}^{3/5}\,n\,, \tag{11.49}$$

where n is the number of neutral hydrogen atoms in 1 cm^3, while $\bar{v}$ is determined by formula (11.48) with $\mu_2 = 1{\cdot}0080$.

In this case it is extremely difficult to obtain the value of C in the laboratory. For this reason the following approximate formula is still used for it:

$$C = \frac{e^2\,\alpha}{h}\,\overline{R_k{}^2}\,. \tag{11.50}$$

Here α is the polarisability of the neutral hydrogen atom, equal to $6{\cdot}63 \times 10^{-25}$ cm^3, while $\overline{R_k^2}$ is the mean square radius of the orbit for the *higher* excited state of the transition $i \rightarrow k$ in question. The value of $\overline{R_k^2}$ has been calculated in a number of cases, of which we shall give the most important. For the *sodium* D lines $\overline{R_k^2} = 41\, a_0^2$, where a_0 is the atomic unit of length (radius of the first Bohr orbit of hydrogen). For the H and K resonance lines of ionised *calcium* $\overline{R_k^2} = 23\, a_0^2$; for the calcium resonance line 4227 Å, $\overline{R_k^2} = 69\, a_0^2$. With the value of $\overline{R_k^2}$ given above for the sodium D lines, for instance, the value of C is found to be $2{\cdot}65 \times 10^{-32}$.

It must be remarked that all the formulae given above for γ_c could be written (as is sometimes done) in terms of the effective cross-sections for broadening. To do so, one can use the formula relating τ_c to the number of collisions Z_c undergone by a given absorbing atom in 1 second:

$$\tau_c = 1/Z_c\,, \tag{11.51}$$

and also the approximate formula which expresses Z_c in terms of the effective cross-section σ for broadening collisions:

$$Z_c = \sigma\, n\, \bar{v}\,. \tag{11.52}$$

Using these formulae, we obtain from (11.9)

$$\gamma_c = 2/\tau_c = 2\, Z_c = 2\, \sigma\, n\, \bar{v}\,. \tag{11.53}$$

Comparing (11.53) with (11.45), (11.47) and (11.49), we can estimate the respective *mean* values of σ also.

We shall discuss later the application of formulae (11.45), (11.47) and (11.49). At the moment we shall consider another "variety" of the pressure effect — the *statistical broadening* of lines.

7. **The statistical broadening of lines.** The clearest example of statistical broadening is that due to the **intermolecular Stark effect**, which arises from the electrostatic fields created by ions and electrons. On account of the presence of these fields, the frequency of the emission (or absorption) of each atom is displaced relative to its normal position. And since for each atom the configuration of the charged particles (ions and electrons) closest to it is in general different, the displacements of the line will also be different for various atoms of the kind in question (at a given moment). The resulting absorption line in the stellar (or solar) spectrum, being formed by a *large* number of atoms with various displacements of the absorption frequency, will therefore be *broadened*.

Consequently, the principal task in this problem is to establish the law which gives the probability $W(F)$ of a given intensity F as a function of the field F itself. In doing so a "standard" field F_0 is usually introduced, to which the field F is referred:

$$F_0 = 2{\cdot}61\, e\, n^{2/3}\,, \tag{11.54}$$

where n is the number of charged perturbing particles in 1 cm^3. The ratio F/F_0 is generally denoted by β. Then the problem reduces to finding the probability $W(\beta)$ that β lies between β and $\beta + \mathrm{d}\beta$. In order to find the probability that the frequency increase lies between ν and $\nu + \mathrm{d}\nu$, it is necessary to know the law which gives the value of the splitting $\Delta\nu$ as a function of the field intensity F. We must distinguish two cases: (1) *the linear Stark effect* (the splitting $\Delta\nu$ is proportional to the perturbing field). This case holds in the broadening of hydrogen lines, and also for a fairly small splitting of some helium lines; (2) *the quadratic Stark effect* ($\Delta\nu$ proportional to F^2). In this case the line splitting is less than for the linear effect. Besides this, the quadratic Stark effect causes the appearance of lines forbidden by the selection rule for the orbital quantum number l.

Let us consider the linear Stark effect for hydrogen lines. Each level of the hydrogen atom (above the first), when in an electric field, undergoes a symmetrical splitting into several components. The number of these components increases with the principal quantum number of the level. The distance of the split levels from the normal position is determined by the formula

$$\Delta E = h\,\Delta\nu = \frac{3}{8\pi^2}\,\frac{h^2}{e m_e}\,F\,n\,n_F\,, \tag{11.55}$$

where the product of the principal quantum number n with the magnetic quantum number n_F is given in Table 6.

Table 6

n	n_F	nn_F	Number of components
1	0	0	1
2	0, ±1	0, ±2	3
3	0, ±1, ±2	0, ±3, ±6	5
4	0, ±1, ±2, ±3	0, ±4, ±8, ±12	7
5	0, ±1, ±2, ±3, ±4	0, ±5, ±10, ±15, ±20	9

The possibility of any transition from a level with principal quantum number n and magnetic number n_F to one with numbers n' and n_F' is determined by the appropriate selection rules. It follows from Table 6 and formula (11.55) that the width of a line of a given series increases with the principal quantum number of the upper level.

It must be noticed that for very strong fields, exceeding 100,000 volts/cm, the Stark effect changes from the linear to the quadratic, and then the line splitting ceases to be symmetrical relative to the unperturbed position. However, such fields apparently do not exist under stellar conditions.

We now turn to the absorption coefficient for the Balmer lines, which are broadened by the *intermolecular* Stark effect. This coefficient is found from the distribution $W(\beta)$, taking account of all permitted transitions between the corresponding levels with $n = 2$ and $n = k$. For the purely *static* case where the perturbing particles (ions and electrons) may be supposed *stationary* with regard to the perturbed hydrogen atom, calculations for the absorption coefficient in the wings of lines of the Balmer series give the formula [168, § 82]

$$s_\lambda = C\, F_0^{3/2}/(\lambda - \lambda_0)^{5/2}\,, \tag{11.56}$$

where for H_α, H_β, H_γ and H_δ the constant C is respectively $3{\cdot}13 \times 10^{-16}$, $0{\cdot}885 \times 10^{-16}$, $0{\cdot}442 \times 10^{-16}$ and $0{\cdot}309 \times 10^{-16}$, $\lambda - \lambda_0$ being expressed in ångströms.

This formula is applicable only in the wings of lines of the Balmer series. A comparison of the theoretical and observed contours of the inner parts of the lines H_α, H_β, H_γ and H_δ often shows a considerable divergence. This is explained chiefly by the fact that formula (11.56) has been deduced for the purely static case. It is quite evident that in reality the perturbing particles are in a state of continuous motion relative to the neutral hydrogen atoms. When this factor is taken into account, it is found that the decrease of s_λ when $\lambda - \lambda_0$ increases in the inner parts of the wings of the lines H_α, H_β, H_γ and H_δ is slower than that given by formula (11.56); further, the manner of dependence of s_λ on λ is greatly changed in the central parts of the lines.

The quadratic Stark effect in the majority of cases appears as *collision damping*, and not as statistical broadening. The same is apparently true for the majority of helium lines (see Chapter 13).

8. **Concluding remarks.** We shall now investigate in which cases we must use the statistical theory, and in which the theory of collision broadening. We begin with broadening by *charged* particles, corresponding to the cases $k = 2$ and $k = 4$.

$k = 2$. The broadening of hydrogen lines and some helium lines *by the linear Stark effect*. It is clear that in this case the index k in formula (11.42) is equal to two, since F is proportional to r^{-2}. A study of the limiting values of $\Delta\lambda_g$, and of the relative part played by the two forms of the pressure effect, shows that in the atmospheres of stars the broadening of the wings of lines of the Balmer series takes place

in accordance with the statistical theory, the broadening particles being apparently only *ions*; it may be, however, that electrons should also be taken into account (C. DE JAGER 1952).

$k = 4$. The *quadratic Stark effect*, due to ions and electrons. The appropriate analysis shows that here, contrary to the preceding case, the broadening agent is damping *by collisions with electrons*. For the case in question we must use formula (11.47), putting $n = n_e$, and calculate $\bar{v}$ for collisions between atoms and *electrons*. For lines where the constant C is fairly large, the form of broadening considered may play an important part in the atmospheres of stars of all spectral classes. The broadening of lines of *ionised elements* by collisions with electrons has not yet been sufficiently studied.

We now turn to line broadening as a result of the interaction of atoms with *neutral particles*.

$k = 3$. Broadening *by self-pressure*. Calculations show that this effect is negligible for metallic lines, by reason of the relatively low content of metals in the atmospheres of stars, i. e. the small values of n in formula (11.45). It may apparently occur in the first few terms of the Balmer series in the spectra of the cooler stars.

$k = 6$. Broadening of lines *by collisions of atoms with neutral hydrogen atoms.* (At temperatures where hydrogen is completely ionised, the broadening may be caused by the next most abundant element, neutral helium.) For $k = 6$, collision damping predominates; the "statistical wings" lie at a great distance from the centre of the line, where $r_\nu \approx 1$. Consequently, in this case we must use formula (11.49) for γ_c. It can also be used in the case where the atoms perturbed by neutral hydrogen are in an ionised state.

We may make the following general remark relating to the use of the formulae given above for γ_c and s_ν. In stars with T_e less than 6000°, the hydrogen is mainly neutral. Hence, on account of the very high content of hydrogen. the broadening of the majority of metal lines takes place through collisions of the atoms with neutral hydrogen [formula (11.49)]. The broadening by collisions with electrons [$k = 4$, formula (11.47)] can take place only for lines which are very sensitive to the quadratic Stark effect (large C). At temperatures above 8000°, the hydrogen is almost wholly ionised. Hence the case $k = 6$ does not hold here, and the broadening of lines takes place through the interaction of the atoms with ions and electrons. For neutral metals, the line broadening is here caused by collisions of the atoms with electrons, and for some helium lines by collisions of helium atoms with protons.

Finally, the broadening of the wings of lines of the Balmer series is determined in the majority of cases by the statistical theory [$k = 2$, for the wings far from the centre, formula (11.56)]. Here the broadening

particles are ions and possibly electrons (see above). In the inner parts of the lines it is necessary to introduce corrections to take account of the motion of the perturbing charged particles relative to the hydrogen atoms. In the atmospheres of supergiants and very hot stars, radiation damping begins to become important.

In conclusion, it must be remarked that, in the cases $k = 4$ and $k = 6$, not only broadening but also displacement of the line occurs. This circumstance must be taken into consideration in studying the gravitational displacement of lines in the Sun.

Chapter 12. The application of the elementary theory of contours. Curves of growth

1. The elementary theory of contours. Let us consider the application to stellar (and solar) absorption lines of the elementary model of a scattering "reversing layer" which we studied in Chapter 9. We recall that for this model we obtained the following formula for the residual intensity:

$$r_\nu = H_\nu/H_\nu^0 = 1/(1 + \tfrac{3}{4}\,\tau_{\nu,\,\sigma}) \,. \tag{12.1}$$

Here we are speaking of the spectrum of the integrated radiation sent out by the whole disc of the star (or of the Sun).

We have noticed in Chapter 9 that the model in question is very schematic and that in reality there is no justification for dividing the atmosphere into the "photosphere" and the "reversing layer". However, by using formula (12.1) it is possible to obtain a satisfactory representation of the contours of many fairly intense absorption lines in solar and stellar spectra. (For faint lines, purely instrumental effects play a considerably greater part than for strong lines.) Moreover, formula (12.1) is also important because by using it we can conduct a large number of investigations of a (so to speak) exploratory nature.

In order to use formula (12.1) to construct an absorption line contour, it is necessary to know s_ν. Except in the case of lines broadened by the Stark effect, we can use formula (11.39) for s_ν. For the case $a \ll 1$, which is the most frequently encountered, the absorption coefficient s_ν can, as we have already said, be represented by means of two expressions, one determined only by the Doppler effect, and the other determined only by damping [see (11.24) and (11.34)].

Since, in formulae (11.24) and (11.34), the absorption coefficient is referred to one atom, we obtain for the optical thickness $\tau_{\nu,\,\sigma}$ in formula (12.1)

$$\tau_{\nu,\,\sigma} = \int n_i\, s_\nu\, \mathrm{d}h \,, \tag{12.2}$$

where n_i is the number of absorbing atoms, of the kind considered, in 1 cm^3 which are in the excitation state i.

In general the coefficient s_ν depends on h. In formula (11.24) δ_{ik} depends on h (through the constant γ_c); in formula (11.34) $\Delta\nu_D$ depends on h. In the former case the dependence on h may be marked; in the latter case it is slight, since T changes little within the "reversing layer", and moreover, T appears in formula (11.33) under the radical sign. However, in order to simplify the problem further, we shall take for δ_{ik} and $\Delta\nu_D$ their *mean* values inside this layer. Then (12.2) can be rewritten

$$\tau_{\nu,\sigma} = s_\nu \int n_i \, \mathrm{d}h = s_\nu N_i \,. \tag{12.3}$$

The number N_i is the number of atoms of the kind considered which are in a column of base 1 cm^2 and height equal to that of the "reversing layer". We shall call N_i simply "the number of absorbing atoms above the photosphere", bearing in mind that it refers to a base area of 1 cm^2.

Having regard to (12.3), the formula (12.1) can be rewritten

$$r_\nu = 1/(1 + \tfrac{3}{4} s_\nu N_i) \,. \quad (12.4$$

Let us now consider what should be the contour of an absorption line when formulae (12.4), (11.24) and (11.34) are applied. We denote by X_0 the optical thickness in the line centre (i. e. for $\nu = \nu_0 \equiv \nu_{ik}$). In accordance with (12.3), (11.34) and (11.40), this optical thickness is*

$$X_0 = \tau_{\nu_0,\sigma} = N_i s_{\nu_0} = \frac{\sqrt{\pi}\, e^2}{m_e c} \frac{f_{ik}}{\Delta\nu_D} N_i \,. \tag{12.5}$$

Thus, in the case of Doppler broadening, the optical thickness $\tau_{\nu,\sigma}$ is

$$s_\nu N_i = \tau_{\nu,\sigma} = X_0 \, e^{-[(\nu-\nu_0)/\Delta\nu_D]^2} \tag{12.6}$$

Let us find the corresponding expression for $\tau_{\nu,\sigma}$ in the case of radiation damping. We notice that damping begins to become important only at a distance $\nu - \nu_{ik}$ from the central frequency $\nu_{ik} = \nu_0$ which is several Doppler widths (see Fig. 25). For the majority of practical cases this frequency interval considerably exceeds δ_{ik} (usually by several orders of magnitude). For this reason we can neglect the second term in the denominator of formula (11.24) for frequencies where radiation damping is important. Thus, using (12.3), (12.5) and (11.24), we obtain

* We recall (see Chapter 11) that, when a does not satisfy the inequality $a \ll 1$, the value of s_{ν_0} determined by formula (11.40) is only approximately the central value of s_ν.

for $\tau_{\nu,\sigma}$, in regions where only the influence of radiation damping is important,

$$s_\nu N_i = \tau_{\nu,\sigma} = \frac{e^2}{m_e c} \frac{\delta_{ik} f_{ik}}{(\nu - \nu_0)^2} N_i = X_0 \frac{\delta_{ik}}{\sqrt{\pi}} \frac{\Delta\nu_D}{(\nu - \nu_0)^2} . \qquad (12.7)$$

Using formulae (12.4), (12.6) and (12.7), we shall consider the change in the form of the contour when N_i increases, and therefore when X_0 increases. So long as the optical thickness of the layer is considerably less than unity ($X_0 \ll 1$), the value of r_ν is very little different from unity, even in the centre of the line, where s_ν is greatest; thus the line is practically invisible on the background of the continuous spectrum. When N_i increases further, a noticeable absorption first occurs in the most central Doppler part of the curve of s_ν. Thus, if for the case represented in Fig. 25, i. e. for the D_1 line at $T = 5700°$, the value of $\tau_{\nu_0,\sigma} = X_0 = 1$, then, even in the transitional part of the curve of s_ν, the optical thickness must be of the order of 10^{-4}, so that at this point r_ν will be practically equal to unity; at the centre of the line, however, r_ν is equal to $(1 + \frac{3}{4})^{-1} = 0.57$.

The part played by damping in the formation of the line contour begins to be noticeable only for N_i such that the value of $\tau_{\nu,\sigma}$ for the transitional connecting part of the coefficient s_ν (see Fig. 25) is no longer so small compared with unity that it can be neglected. Before this critical value of N_i is reached, the absorption line contour is determined by the Doppler effect, i. e. by formula (12.6). Fig. 26 shows contours calculated from formulae (12.4), (12.5) and (12.6) for $\Delta\lambda_D = 0.01$ Å and various values of the product* $N_i f_{ik}$. We notice that the absorbing effect of the atoms is determined in formulae (12.6) and (12.7) not by the number N_i itself, but by the *product* $N_i f_{ik}$.

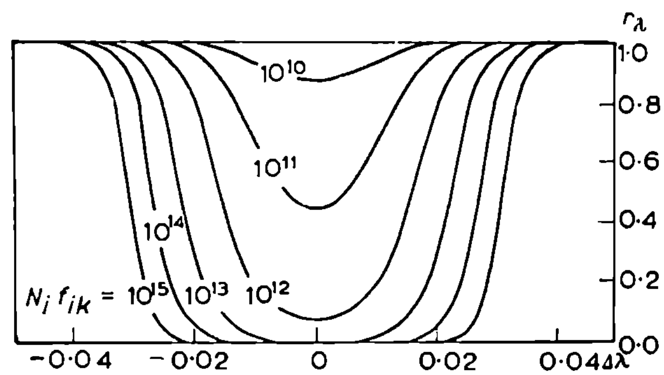

FIG. 26

An examination of Fig. 26 reveals the very characteristic form of contours due to the Doppler effect: the rapid decrease of r_ν towards the line centre (for fairly large $N_i f_{ik}$).

When $N_i f_{ik}$ increases further, the parts of the curve of s_ν which are determined by radiation *damping* become more and more effective. If the Doppler effect were completely absent, the line contours would

* Figures 26, 27, 28 and 30 are taken from the review article by O. A. MEL'NIKOV [81]. In constructing the contours, the factor $\frac{3}{4}$ in the denominator of (12.4) was replaced by 1.

have the form shown in Fig. 27, where the value of δ_{ik} has been transformed to the wavelength scale by means of (4.38), and is 0·001 Å. This value is approximately an order of magnitude larger than the natural width, but, as we shall see later, the constant γ_c for collision damping is often an order of magnitude larger than γ_0 (in particular, for lines of metals in the Sun's atmosphere).

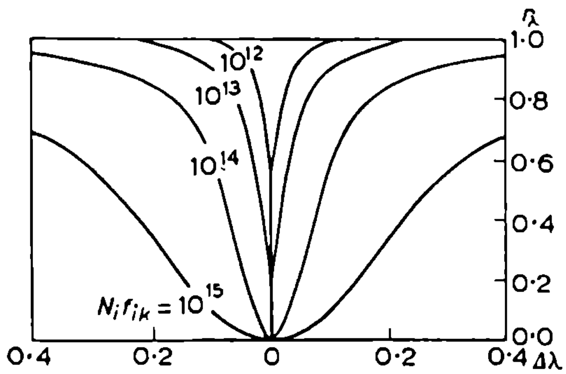

Fig. 27

We see from Fig. 27 that, for pure damping (contrary to the case where the contour is determined by the Doppler effect), very well-defined wings are formed in the lines, and these pass very smoothly into the continuous spectrum.

In Fig. 28 we give contours constructed from formulae (12.4) and (11.39), using the appropriate tables for s_ν. The initial values were: $\Delta\lambda_D = 0{\cdot}1$ Å; δ_{ik} (transformed to the wavelength scale) $= 0{\cdot}001$ Å.

It follows from Fig. 28 that for relatively small $N_i f_{ik}$ the absorption line contour is determined mainly by the Doppler effect, but for large $N_i f_{ik}$ by *damping*. It is true that, even for large $N_i f_{ik}$, the variation of the intensity at the very centre of the line is determined by the

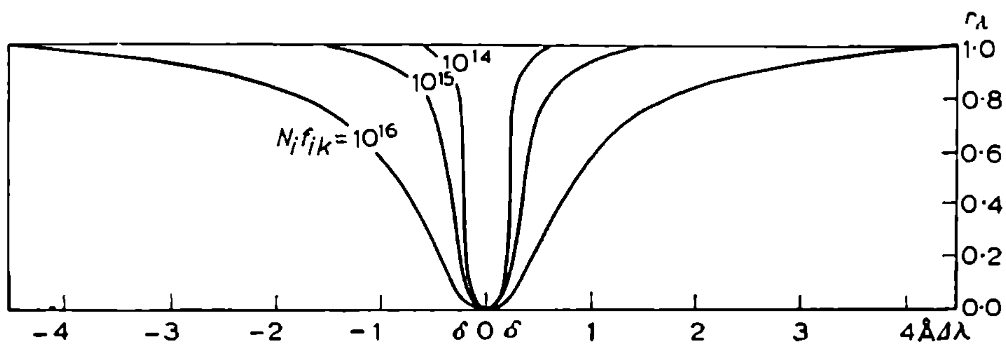

Fig. 28

Doppler effect, but for the frequencies concerned (δ — δ in Fig. 28) r_ν is so small that it is immaterial to the observer which effect plays the chief part in this frequency interval. Hence we can say that, for *large* $N_i f_{ik}$, the *whole* contour is determined mainly by *damping*.

Consequently, the contours of *faint* absorption line (where there are no marked wings) are determined mainly *by the Doppler effect*, while for fairly *strong* lines (with well-developed wings) they are determined mainly *by damping effects*.

2. Comparison of theoretical and observed contours. Let us consider to what extent the theoretical contours agree with those obtained from observation. Such a comparison is the most reliable for the stronger absorption lines. The contours of faint lines are usually very much distorted by purely instrumental effects. The removal of these distorting effects is not generally sufficiently trustworthy. For strong lines, the contour is determined mainly by damping processes. Introducing (12.7) into (12.4), we obtain the formula

$$r_\nu = \frac{1}{1 + \frac{3}{4}\frac{e^2}{m_e c}\frac{\delta_{ik}}{(\nu-\nu_0)^2} N_i f_{ik}}. \tag{12.8}$$

The comparison of the theoretical contour, given by formula (12.8), with the observed contour is carried out as follows. By inserting in (12.8) a number of values for the product $\delta_{ik} f_{ik} N_i$, we attempt to obtain the best possible agreement between theory and observation. [The same has to be done for $r_\nu(\theta)$ in formula (9.24).] It is found that this agreement is satisfactory for the contours of many strong lines in solar and stellar spectra. Fig. 29 shows an example of this, for the H and K lines in the spectrum of the centre of the Sun's disc.

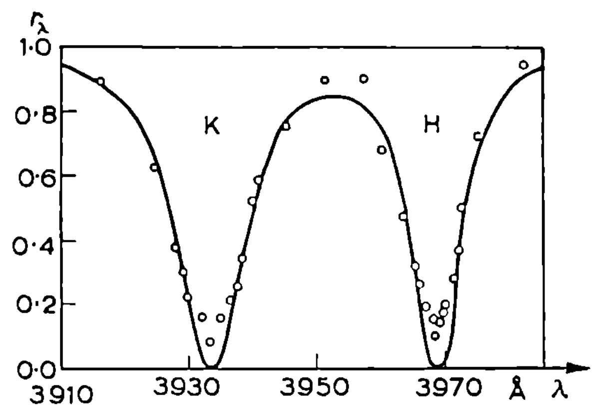

Fig. 29

It is seen from Fig. 29 that theory and observation are in greatest disagreement in the *central* parts of the line. The observed central intensity is greater than the theoretical. This excess of the observed central intensity over the theoretical exists, in fact, for all strong absorption lines. Despite this divergence (to which we shall return in Chapter 14), Fig. 29 and similar figures indicate that damping is the principal factor which determines the contours of strong absorption lines. This conclusion is confirmed by numerous other data. A similar result also follows from a theory which is much more rigorous than that leading to the original formula (12.1).

Taking these results into account, we can use the values of $N_i f_{ik} \delta_{ik}$, found in the manner described above, to determine the number of atoms above 1 cm² of the photosphere, i. e. to determine the value of N_i. To do so, we must know f_{ik} and δ_{ik}. The values of $f_{ik} = f_{1k}$ for resonance lines are usually known. As a rule they are of the order of unity. The values of $\delta_{ik} = \delta_{1k}$ are more uncertain. According to (11.25), they are

determined not only by radiation damping, but also by collision damping. The value of γ_c is known as yet only in a small number of cases (and then only approximately). On the average, for lines of the *solar spectrum*, γ_c is approximately 5 to 10 times greater than γ_0 (the classical damping constant). If we omit the collision damping constant in δ_{ik}, we shall *overestimate* the number of atoms obtained, and, as follows from formula (12.8), we shall make an error in the value of N_i amounting to a factor of the order of $\sqrt{[(\gamma_i + \gamma_k)/\gamma_c]}$. If we take into account that for resonance lines $\gamma_i + \gamma_k \approx \gamma_k \approx \gamma_0$ and use the maximum factor 10, the value of N_i found by omitting the constant γ_c is exaggerated by a factor of about 3.

We can find the value of N_i in the manner described, not only from formula (12.8), which refers to the whole disc of the Sun (or star), but also from formula (9.24), which is applicable to *individual* points on the solar disc. The value of $\tau_{\nu,\sigma}$ which appears in this formula must again be taken from the expression (12.7).

The values of N_1 obtained by A. Unsöld for resonance lines in the solar spectrum (in the part of the spectrum accessible to direct observation) are given in Table 49 in the first (1938) edition of his *Physik der Sternatmosphären* [168]. These values are overestimated, since collision damping was not taken into consideration in obtaining them.

Similar calculations can be used for the study of the chemical composition of stellar atmospheres. Here, using the formulae of Boltzmann and Saha, we can go from the number of atoms in some one state to the number in other states and, by summing, find the total number of atoms of the element considered. However, besides the inadequacy of the "reversing layer" model which is used, this method of determining the number of atoms suffers from the defect of being inapplicable for faint lines, since their contours are almost entirely of instrumental origin. Because of this circumstance, a method of determining the values of N_i not from line contours but *from the total absorption*, i. e. from the equivalent width of the line, has great advantages. The instrumental distortion has very much less effect on the equivalent width than on the contour itself. The chief distorting factor in this case is scattered light, and ghosts in the case of a grating spectrograph.

3. **The elementary theory of curves of growth.** A curve which relates the equivalent widths of lines (i. e. the quantities W_λ or W_ν) to the values of $f_{ik} N_i$ is called a **curve of growth.** In constructing a theoretical curve of growth, we again use formula (12.4). However, it must be borne in mind that, even for the "reversing layer" scheme considered, formula (12.4) has been deduced by making several mathematical simplifications. A more rigorous mathematical analysis shows that, instead of the factor $\frac{3}{4}$ in the denominators of (12.1) and (12.4), $\tau_{\nu,\sigma}$ should be preceded by some factor μ which itself depends on $\tau_{\nu,\sigma}$ and

is unity for $\tau_{\nu,\sigma} = 0$, while for $\tau_{\nu,\sigma} \to \infty$ it tends to $\frac{3}{4}$. Bearing in mind the approximate nature of our calculations, we can put this factor $\mu(\tau_{\nu,\sigma})$ equal to unity. For this reason we shall use, instead of formula (12.4), the even simpler formula

$$r_\nu = 1/(1 + \tau_{\nu,\sigma}) = 1/(1 + N_i s_\nu) . \tag{12.9}$$

If, instead of assuming $\mu = 1$, we introduce μ explicitly, this factor will appear in the final formulae for the equivalent width as a factor in front of N_i, and its effect is thus that the values of N_i which we determine will differ from the true values by a factor μ.

We obtain from (12.9) for the depth R_ν

$$R_\nu = 1 - r_\nu = \frac{\tau_\nu}{1 + \tau_\nu} = \frac{N_i s_\nu}{1 + N_i s_\nu} = \frac{1}{1 + (N_i s_\nu)^{-1}} . \tag{12.10}$$

Consequently, we have for the equivalent width W_ν, in accordance with (9.2),

$$W_\nu = \int_0^\infty \frac{d\nu}{1 + (N_i s_\nu)^{-1}} = 2 \int_{\nu_0}^\infty \frac{d\nu}{1 + (N_i s_\nu)^{-1}} . \tag{12.11}$$

The last member of equation (12.11) is obtained by using the symmetry of the function s_ν about the central frequency ν_0 [see (12.6) and (12.7)]. Consequently, according to (12.10), R_ν is also symmetrical, and both halves of the contour give the *same* total absorption.

The integration in (12.11) extends, strictly speaking, over the whole spectrum, but in practice it is restricted to the relatively narrow region occupied by the absorption line for which W_ν is calculated, since s_ν becomes practically equal to zero even for relatively small values of $\nu - \nu_0$.

Observers generally use values of W_λ and not W_ν. Further, as we shall see, it is convenient to use not the values of W_λ itself, but those of W_λ/λ. Because of the relatively small values of equivalent widths, we can use for the connection between W_λ and W_ν the differential formula (4.38), and this gives

$$W_\lambda/\lambda = W_\nu/\nu . \tag{12.12}$$

Let us now investigate the dependence of W_ν, and so of W_λ, on X_0, i. e. on the values of $N_i f_{ik}$. We shall first consider values of $N_i f_{ik}$ for which the line widths are determined only by the Doppler effect. In this case we can use for $N_i s_\nu$ the expression (12.6); using (12.11) we find

$$W_\nu = 2 \int_{\nu_0}^\infty \frac{d\nu}{1 + e^{[(\nu-\nu_0)/\Delta\nu_D]^2}/X_0} . \tag{12.13}$$

We now introduce a new variable p:

$$p = (\nu - \nu_0)/\Delta\nu_D \,. \tag{12.14}$$

Then (12.13) takes the form

$$W_\nu = 2\,\Delta\nu_D \int_0^\infty \frac{\mathrm{d}p}{1 + e^{p^2}/X_0} = 2\,X_0 \Delta\nu_D \int_0^\infty (e^{p^2} + X_0)^{-1}\,\mathrm{d}p \,. \tag{12.15}$$

We first consider the case where $N_i f_{ik}$ is so small that $X_0 < 1$. In this case the integral (12.15) can be calculated by means of a binomial expansion of the integrand:

$$W_\nu = 2\,\Delta\nu_D\,X_0 \int_0^\infty (e^{-p^2} - X_0\,e^{-2p^2} + X_0^2\,e^{-3p^2} - \dots)\,\mathrm{d}p \,. \tag{12.16}$$

Since

$$\int_0^\infty e^{-np^2}\,\mathrm{d}p = \tfrac{1}{2}\,\sqrt{(\pi/n)} \,, \tag{12.17}$$

we find

$$W_\nu = \sqrt{\pi}\,\Delta\nu_D\,X_0 \left\{1 - \frac{X_0}{\sqrt{2}} + \frac{X_0^2}{\sqrt{3}} - \dots\right\} . \tag{12.18}$$

This expansion is applicable in practice for $0 \leq X < 0{\cdot}5$.

To calculate (12.15) when X_0 increases further, we put

$$X_0 = e^b \,,\; b = \log_e X_0 ,\; p^2 = u \,. \tag{12.19}$$

In this case (12.15) takes the form

$$W_\nu = \Delta\nu_D \int_0^\infty \frac{u^{-\frac{1}{2}}\,\mathrm{d}u}{1 + e^{u-b}} \,. \tag{12.20}$$

The integral on the right-hand side of (12.20) is of great importance in the electron theory of metals. For fairly large $b = \log_e X_0$, it can be represented by the following asymptotic expression:

$$\int_0^\infty \frac{u^{-\frac{1}{2}}\,\mathrm{d}u}{1 + e^{u-b}} = 2\,\sqrt{b}\left\{1 - \frac{\pi^2}{24\,b^2} - \frac{7\,\pi^4}{384\,b^4} - \dots\right\} . \tag{12.21}$$

In view of (12.21) and (12.19), we can rewrite (12.20) as

$$W_\nu = 2\,\Delta\nu_D\,\sqrt{(\log_e X_0)}\left\{1 - \frac{\pi^2}{24\,\log_e^2 X_0} - \frac{7\,\pi^4}{384\,\log_e^4 X_0} - \dots\right\} . \tag{12.22}$$

This series can be used in practice for $\log_e X_0 \geq 4$ or $X_0 \geq 55$. For $0{\cdot}5 < X_0 < 55$, the integral (12.20) or its equivalent (12.15) must be evaluated by methods of numerical integration.

When $N_i f_{ik}$ (and consequently X_0) increases further, the part played by the Doppler effect in producing the line contour becomes entirely negligible, and the line contour is determined by damping processes. For very large values of X_0, we introduce (12.7) into (12.11):

$$W_\nu = 2 \int\limits_{\nu_0}^{\infty} \frac{d\nu}{1 + \sqrt{\pi}(\nu - \nu_0)^2/X_0\,\delta_{ik}\,\Delta\nu_D} \,. \tag{12.23}$$

Putting

$$\sqrt{\pi}(\nu - \nu_0)^2/X_0\,\delta_{ik}\,\Delta\nu_D = x^2 \,, \tag{12.24}$$

we find

$$W_\nu = 2\,\pi^{-\frac{1}{4}} \sqrt{(X_0\,\delta_{ik}\,\Delta\nu_D)} \int\limits_0^{\infty} \frac{dx}{1 + x^2} = \pi^{\frac{3}{4}} \sqrt{(X_0\,\delta_{ik}\,\Delta\nu_D)} \,. \tag{12.25}$$

The intermediate part of the curve of growth, corresponding to the transition from (12.22) to (12.25), is found by methods of numerical integration from the original formula (12.11), using the values of s_ν determined, for example, from the tables in Appendix I.

Transforming now to the wavelength scale by formula (12.12), and using the expression (11.31) for $\Delta\nu_D$, we obtain for the three cases we have considered:

$0 \leq X_0 < 0{\cdot}5$:

$$\frac{W_\lambda}{\lambda} = \sqrt{\pi}\,\frac{v_0}{c}\,X_0 \left\{1 - \frac{X_0}{\sqrt{2}} + \frac{X_0^2}{\sqrt{3}} - \ldots\right\} \approx \frac{v_0}{c}\sqrt{\pi}\,X_0 \,, \tag{12.26}$$

$X_0 \geq 55$:

$$\begin{aligned}\frac{W_\lambda}{\lambda} &= 2\,\frac{v_0}{c}\sqrt{(\log_e X_0)} \left\{1 - \frac{\pi^2}{24\log_e^2 X_0} - \right. \\ &\left. - \frac{7\pi^4}{384\log_e^4 X_0} - \ldots\right\} \approx 2\,\frac{v_0}{c}\sqrt{(\log_e X_0)} \,,\end{aligned} \tag{12.27}$$

X_0 very large:

$$\frac{W_\lambda}{\lambda} = \pi^{\frac{3}{4}} \sqrt{\left(\frac{v_0}{c}\,\frac{\delta_{ik}}{\nu_0}\,X_0\right)} = \frac{\pi^{\frac{1}{4}}}{2} \sqrt{\left(\frac{v_0}{c}\,X_0\,\frac{\Gamma_{ik}}{\nu_0}\right)} \,. \tag{12.28}$$

It follows from these formulae that, to remove the dependence of the curve of growth on the wavelength of the lines considered, it is in fact necessary to use not W_λ but W_λ/λ. The frequency ν_0 (and therefore λ_0) appears only in (12.28), and then under the radical sign.

The following conclusions can be drawn from these formulae. For faint lines, where $X_0 \ll 1$, the equivalent width increases proportionally to X_0, i. e. proportionally to the number $N_i f_{ik}$. This relatively rapid increase of W_λ is here due to the fact that when X_0, i. e. $N_i f_{ik}$, increases, a "saturation" of the line takes place in its central parts, so that r_ν diminishes from unity (when $X_0 = 0$) almost to zero (when $\log_e X_0 \gg 1$), and the depth in the line centre increases from 0 to approximately 1. When X_0 increases further, and $r_{\nu_0} \ll 1$, the increase of W_λ takes place not because of an increase of the *depth* R_ν in the central parts of the line, but mainly because of an increase in its width. It follows from an analysis of formulae (12.9) and (12.6) that the change in the line width when X_0, i. e. $N_i f_{ik}$, increases, is extremely slow, because of the presence of the exponential factor in formula (12.6). Finally, when X_0 is already very large and increases still further, the equivalent width increases proportionally to the square root of $N_i f_{ik}$.

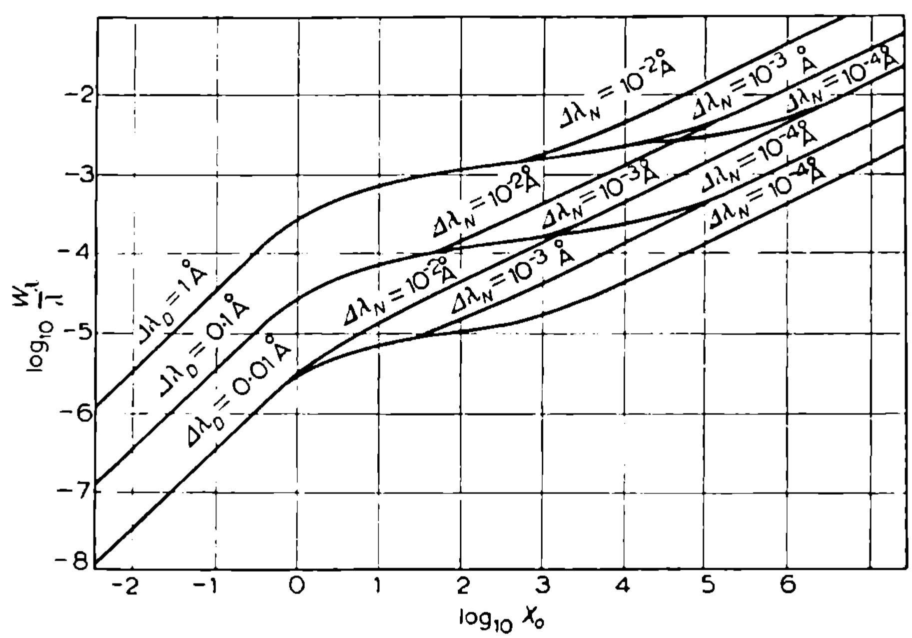

FIG. 30

For a number of reasons, the curve of growth is constructed, not directly as a relation between W_λ/λ and X_0, but as a relation between the (common) logarithms of these quantities. Fig. 30 shows a curve of growth for various values of $\Delta\lambda_D$ and δ_{ik}, constructed from the formulae (12.26) to (12.28) above. We denote by $\Delta\lambda_N$ the value of δ_{ik} expressed in the wavelength scale. According to (4.38), we have

$$\Delta\lambda_N = (\lambda_0^2/c)\,\delta_{ik}\,. \tag{12.29}$$

It follows from Fig. 30 that, when $\Delta\lambda_D$, i. e. the velocity dispersion v_0, increases, the entire curve of growth is raised, while the first two portions

are raised more than the third. Also, for a given $\Delta\lambda_D$ the third and straightest part of the curve of growth, which corresponds to damping, is raised when $\Delta\lambda_N$, i. e. δ_{ik}, increases. Thus the curves of growth in Fig. 30 form a family depending on the two parameters $\Delta\lambda_D$ and $\Delta\lambda_N$.

Theoretical curves of growth are often constructed so that not $\log_{10}(W_\lambda/\lambda)$ but $\log_{10}(W_\lambda c/\lambda v_0)$ appears on the axis of ordinates. In this case formulae (12.26) — (12.28) take the form

$$\frac{W_\lambda}{\lambda}\frac{c}{v_0} = \frac{W_\lambda}{\Delta\lambda_D} = \sqrt{\pi}\, X_0 \left\{1 - \frac{X_0}{\sqrt{2}} + \frac{X_0^2}{\sqrt{3}} - \dots\right\} \approx \sqrt{\pi}\, X_0\,, \quad (12.30)$$

$$\frac{W_\lambda}{\lambda}\frac{c}{v_0} = \frac{W_\lambda}{\Delta\lambda_D} = 2\sqrt{(\log_e X_0)}\left\{1 - \frac{\pi^2}{24\log_e^2 X_0} - \frac{7\pi^4}{384\log_e^4 X_0} - \dots\right\} \approx 2\sqrt{(\log_e X_0)}\,, \quad (12.31)$$

$$\frac{W_\lambda}{\lambda}\frac{c}{v_0} = \frac{W_\lambda}{\Delta\lambda_D} = \frac{\pi^{\frac{1}{4}}}{2}\sqrt{(X_0 Z)} = \pi^{\frac{3}{4}}\sqrt{(a X_0)}\,, \quad (12.32)$$

where

$$Z = \frac{\Gamma_{ik}}{\Delta\nu_D} = \frac{\Gamma_{ik}}{\nu_0}\frac{c}{v_0} = 4\,\pi\,\frac{\delta_{ik}}{\Delta\nu_D} = 4\,\pi\, a\,. \quad (12.33)$$

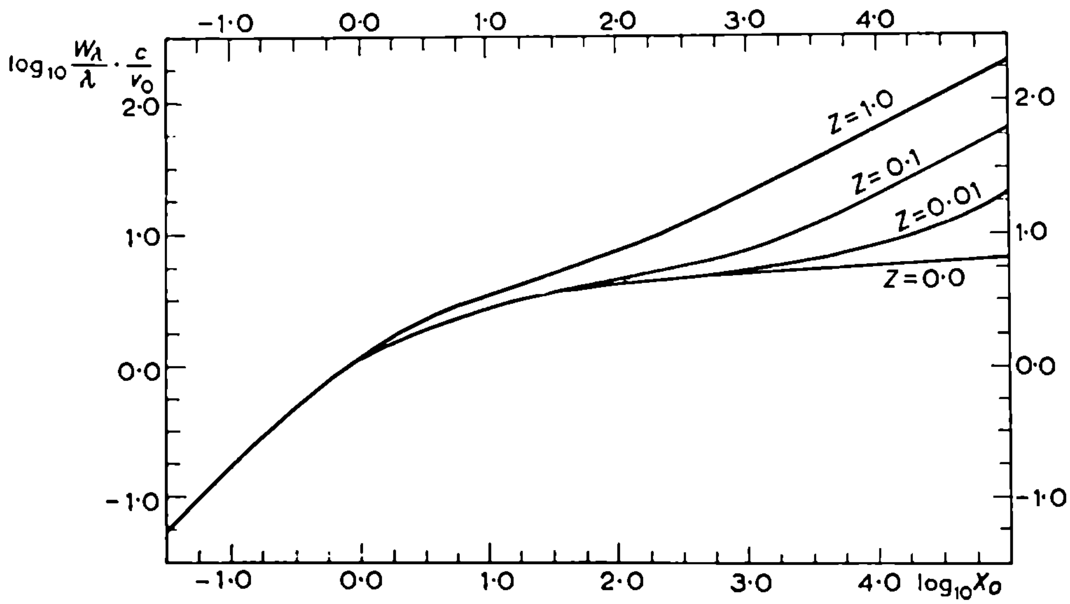

Fig. 31

The value of a is determined by formula (11.38), and $Z = 4\,\pi\,a$.

In the case considered, the ordinate is a function of X_0 alone for the first two parts of the curve of growth; the last portion is determined also by the ratio of δ_{ik} to $\Delta\nu_D$. This is all shown graphically in Fig. 31. We see that, in this case, only a branching of the curve of growth takes place, corresponding to different values of Z, and therefore of a.

4. Methods of constructing curves of growth from observations. Curves of growth can be constructed from observations only by using lines which appear in multiplets. It follows from formula (12.5) that $\log_{10} X_0$, i. e. the abscissa of the curve of growth, is equal, for a given element, to $C + \log_{10}(N_i f_{ik}) = C + \log_{10} N_i + \log_{10} f_{ik}$, where C is independent of N_i and f_{ik}. It thus follows that N_i and f_{ik} are on an exactly equal footing. The construction of an empirical curve of growth from observation must therefore take place as follows. We select some multiplet whose components have a common lower level. For such a multiplet, N_i is practically the same for all the lines composing it. (There is a small difference in N_i connected with the fact that such a level is, as a rule, composite, the sub-levels differing among themselves by some small amount, from some hundredths of an electron-volt to a few tenths of an electron-volt.) We assume also that the relative values (or, better still, the absolute values) of f_{ik} are known for all lines of the given multiplet. We can then represent this multiplet in the following way. On the axis of ordinates in Fig. 31, we mark off the values of $\log_{10}(W_\lambda/\lambda)$ found from observation for each line of the multiplet, and on the axis of abscissae we mark off the corresponding $\log_{10} f_{ik}$, marking on the graph the corresponding points. We then join all the points thus obtained by a smooth line. Consequently, one multiplet gives us a segment of the required curve of growth. (In the majority of cases the range of values of f_{ik} in one multiplet is insufficient to construct the entire curve of growth.)

We must carry out the same procedure for other multiplets also. From the separate portions of the curve, found in the manner described, we can construct the whole curve of growth by displacing the portions of the curve parallel to the axis of abscissae and aiming at a minimum scatter of points about the final curve. This method is very similar to the construction of the characteristic curve of a photographic plate in photographic photometry.

We move the curve of growth, constructed by this method, on Fig. 31 upwards and sideways as a whole, in order to make it coincide as well as possible with one of the theoretical curves shown in Fig. 31. In doing this, the upward movement of the constructed curve (to coincidence with the selected theoretical curve) gives directly $\log_{10}(c/v_0)$, i. e. the velocity v_0. Also, this comparison fixes the value of Z, i. e. the constant Γ_{ik}/ν_0, v_0 being already determined. Finally, by bringing into coincidence the theoretical curve and that constructed from observation, we can determine $\log_{10} X_0$ for any line from the $\log_{10}(W_\lambda/\lambda)$ known from observation. If the absolute value of f_{ik} also is known for this line, then we can determine N_i, by means of formula (12.5), from the known X_0 and $\Delta\nu_D$. (This is another method of determining the relative chemical composition of stellar atmospheres.)

We notice that, in constructing the curve of growth by the above method, and in comparing it with the family of theoretical curves, it is supposed, firstly, that Γ_{ik}/v_0 is the same for all the lines and, secondly, that v_0 is the same for all the elements whose lines are used in constructing the curve of growth.

Let us now consider another method of constructing the curve of growth, which is very similar in principle to the foregoing, but differs from it in that the excitation temperature of the atoms can be obtained. Let N_r be the total number of atoms of the kind considered, in the rth ionisation state (above 1 cm^2 of the "photosphere"). Then, assuming Boltzmann's Law to be applicable, we can write, from (5.54), for the number N_i (the number of atoms in the ith excited state)

$$N_i \equiv N_{r,i} = (N_r/u_r)\, g_i\, e^{-\varepsilon_i/kT_{\text{ex}}}\,, \tag{12.34}$$

where T_{ex} is a temperature which we shall call the **excitation temperature.** (The suffix r has been omitted from the statistical weight and the excitation potential.) Introducing (12.34) into (12.5) and using (11.31), we obtain

$$\begin{aligned}\log_{10} X_0 &= \log_{10}\left\{\frac{\sqrt{\pi}\, e^2}{m_e c}\,\frac{N_r}{u_r}\,\frac{1}{v_0}\right\} + \log_{10}(g_i f_{ik}\,\lambda_0) - \frac{\varepsilon_i}{kT_{\text{ex}}}\log_{10} e \\ &= \log_{10}\left\{\frac{\sqrt{\pi}\, e^2}{m_e c}\,\frac{N_r}{u_r}\,\frac{1}{v_0}\right\} + \log_{10}(g_i f_{ik}\,\lambda_0) - \frac{5040}{T_{\text{ex}}}\,\varepsilon_i\,,\end{aligned} \tag{12.35}$$

where in the last member of this equation ε_i is expressed not in CGS electrostatic units, but in electron-volts.

We now assume that we know either the relative or the absolute values of f_{ik} for all the lines forming the multiplets of the given r times ionised element*. Giving T_{ex} a provisional value, which is generally somewhat less than the effective temperature of the star (see below), we mark off on the axis of abscissae the quantity

$$\log_{10} X_f = \log_{10}(g_i f_{ik}\,\lambda_0) - (5040/T_{\text{ex}})\,\varepsilon_i\,, \tag{12.36}$$

and on the axis of ordinates, $\log_{10}(W_\lambda/\lambda)$. As a result, we obtain a curve of growth for the given r times ionised element, and this curve is to be improved by making T_{ex} more accurate. Then, by seeking the best coincidence (as above) of this curve with one of the theoretical curves in Fig. 31, we obtain the values of v_0 and Γ_{ik}/v_0.

The value of the excitation temperature can be refined as follows. We rewrite (12.35) as

$$Y = \log_{10} X_0 - \log_{10}(g_i f_{ik}\,\lambda_0) = L - (5040/T_{\text{ex}})\,\varepsilon_i\,, \tag{12.37}$$

* The relative values must be on the same scale for all the multiplets, so that, to pass from the relative values of f_{ik} to the absolute values, the former are multiplied by the same factor for all the lines of all the multiplets.

where

$$L = \log_{10} \left\{ \frac{\sqrt{\pi}\, e^2}{m_e c} \frac{N_r}{u_r} \frac{1}{v_0} \right\} . \tag{12.38}$$

Then we mark, on the graph which has already been used to find the coincidence of the theoretical curve of growth and the preliminary curve of growth obtained from observation (i. e. the dependence of $\log_{10}(W_\lambda c/\lambda v_0)$ on $\log_{10} X_0$), the multiplets which served for the construction of this graph. As abscissa for each line we take the corresponding value of $\log_{10} (g_i f_{ik} \lambda_0)$, and as ordinate, the value of $\log_{10}(W_\lambda c/\lambda v_0)$, where W_λ is the observed value of the equivalent width, and v_0 is the velocity, which we have already determined in the first approximation. Then the horizontal distance from the given multiplet to the curve of growth is just the quantity

$$Y = \log_{10} X_0 - \log_{10} (g_i f_{ik} \lambda_0) .$$

In this way we can determine Y for all the multiplets which appear in the curve of growth.

If the assumption which we have made regarding the Boltzmann distribution of atoms is justified, the relation between the values of Y and ε_i found for different multiplets should, according to (12.37), be linear. Then, placing the values of ε_i on the axis of abscissae and those of Y on the axis of ordinates, we obtain a straight line from whose slope we can determine T_{ex}. By producing this line to the point where $\varepsilon_i \equiv 0$ we find L, and thereby N_r, v_0 being already determined. This also is a method of chemical analysis of stellar atmospheres. Of course, it is necessary in the analysis to take account of the other ionisation states of the element considered. (It must be remarked that we could find the excitation temperature T_{ex} without appealing to the theoretical curve of growth, by using only the equation (12.36). In this case there would be a linear relation between $Y' = \log_{10} X_f - \log_{10} (g_i f_{ik} \lambda_0)$ and the excitation potential ε_i. The defect of this method is that it does not give the constant L.)

The linear dependence mentioned is found most accurately by using the method of least squares. Further, the more exact value of T_{ex} obtained can be used once more to construct a curve of growth with the argument $\log_{10} X_f$, etc. The values of v_0 and Γ_{ik}/v_0 are found more accurately at the same time.

Curves of growth may be similarly found for the other ionisation states of the element considered, and for other elements. All these curves of growth can be reduced to a common curve, though, as we shall see later, the curves of growth for neutral and ionised elements are sometimes found to be different.

The curves of growth obtained for different elements may also be different. Hence, strictly speaking, the curve of growth should be constructed separately for each element (in each ionisation state). The comparison of such curves leads to important results (see below).

5. **Laboratory, theoretical, and "solar" line intensities.** Before going on to discuss the principal results obtained from curves of growth, we must also briefly consider the topic of the values of f_{ik} used for the curve of growth [81]. The most reliable values of f_{ik} for multiplets are at present those found in the laboratory. These "intensities" are always given in the form of the product $g_i f_{ik}$, and are determined in the laboratory either from absorption lines or from emission lines. The accuracy of the **laboratory relative intensities** $g_i f_{ik}$ is usually about 10 % on the average. The absolute values of $g_i f_{ik}$ are much less accurate in many cases.

Since laboratory data on the values of $g_i f_{ik}$ exist as yet only for a few elements (chiefly Fe I, Ti I, Ni I, V I), **theoretical intensities** of multiplet lines are also used for the construction of curves of growth. Let us consider, for instance, the group of all possible transitions between some two electron configurations. Such transitions are divided into several multiplets. Let S be the relative integrated intensity of some multiplet (with respect to the other multiplets of the group of transitions considered), and s the relative intensity of a given line in this multiplet. Then in the total of all the transitions, the relative intensity of this line is

$$S\, s/\Sigma\, s\,, \tag{12.39}$$

where the sum is extended over all the lines of the given multiplet, and the quantities $S\, s/\Sigma\, s$ for the given group of transitions (between two electron configurations) are proportional to the quantities $g_i f_{ik} \lambda_0$ [81].

Since there are usually not enough multiplets belonging to one group of transitions for a reliable construction of the curve of growth, several such groups are used. A separate curve of growth is constructed for each of them, and a resultant curve is constructed from these by parallel displacement along the axis of abscissae (as explained above, since $S\, s/\Sigma\, s$ for each class of transitions is proportional to $g_i f_{ik} \lambda_0$).

The calculations of the quantities $S\, s/\Sigma\, s$ have been based hitherto on the assumption that Russell-Saunders coupling can justifiably be applied to all atoms [59, p. 128]. However, it is known that this assumption is erroneous in many cases. A comparison shows that in some cases the laboratory intensities $g_i f_{ik} \lambda_0$ may differ from the theoretical $S\, s/\Sigma\, s$ by a factor of 100, although the laboratory values of $g_i f_{ik} \lambda_0$ generally contain errors of no more than 10 %. Hence **solar line intensities**, as they are called, are often used; these are obtained as follows.

We assume that we have constructed a curve of growth for the Sun from the most reliable existing laboratory values of $g_i f_{ik}$ (say, for Fe I or Ti I). Here the values of $\log_{10} X_f$ are placed on the axis of abscissae, and those of $\log_{10}(W_\lambda/\lambda)$ on the axis of ordinates. The excitation temperature $T_{\rm ex}$ for the Sun is determined as above (from the same lines). Then, for any line in the solar spectrum whose $g_i f_{ik}$ value is unknown, we can find the "solar" value of $\log_{10} X_i$, from the "standard" curve just mentioned, corresponding to the measured W_λ for this line.

We now distribute, according to elements and ionisation states, all the lines of the stellar or solar spectrum whose $g_i f_{ik}$ values are unknown; lines belonging to a given element in a given ionisation state are arranged in groups according to the excitation potential ε_i of the lower level, for example: from 0 to 0·5 eV, from 0·5 to 1·0 eV, etc. The variation of $\log_{10}(g_i f_{ik} \lambda_0)$ for each such group is then determined by the variation of $\log_{10} X_f$, in accordance with (12.36). For each line of such a group $\log_{10}(W_\lambda/\lambda)$, obtained from measurements, is placed according to the "solar" value of $\log_{10} X_f$ for it, and a smooth curve is drawn through all the points of the group, giving a segment of the curve of growth. All such separate segments are moved, as described above, in a horizontal direction, and we obtain the resultant curve of growth. After doing this, we find the best possible coincidence of this curve with a theoretical curve of growth. We do the same with our original "standard" curve of growth constructed from the most reliable $g_i f_{ik}$. In each case, we can find $\log_{10} X_0$ (stellar and solar respectively) from the measured W_λ/λ and the v_0 already determined, and consequently the difference $(\log_{10} X_0)_{\rm st} - (\log_{10} X_0)_\odot$.

The excitation temperature for a given element can be determined as follows. Writing formula (12.35) once for the Sun and again for the star and subtracting, we obtain

$$(\log_{10} X_0)_{\rm st} - (\log_{10} X_0)_\odot = \log_{10} \frac{(N_r)_{\rm st}}{(N_r)_\odot} \frac{(u_r)_\odot}{(u_r)_{\rm st}} - \log_{10} \frac{(v_0)_{\rm st}}{(v_0)_\odot} - 5040\, \varepsilon_i \left[\frac{1}{(T_{\rm ex})_{\rm st}} - \frac{1}{(T_{\rm ex})_\odot}\right] . \tag{12.40}$$

It follows from this that, for the lines of any atom or ion, the difference in the values of $\log_{10} X_0$ is a *linear* function of ε_i. Consequently, on marking off on the axis of abscissae the difference $(\log_{10} X_0)_{\rm st} - (\log_{10} X_0)_\odot$ for various groups of lines, and on the axis of ordinates the potential ε_i, we should obtain a straight line. The gradient of this line determines the value of $[1/(T_{\rm ex})_{\rm st} - 1/(T_{\rm ex})_\odot]$ and, if $(T_{\rm ex})_\odot$ is known, then so is $(T_{\rm ex})_{\rm st}$. This method of determining $(T_{\rm ex})_{\rm st}$ is also used in the case where the difference in $\log_{10} X_f$ is found, and not that in $\log_{10} X_0$. The formula (12.40), which is related to the adjustment of the observed curve to the theoretical, is important, because by means of it we can

study the chemical composition, compared with that of the Sun, of stars close to the Sun in spectral type.

Finally, it should be mentioned that the method described for finding $(T_{ex})_{st}$ from $(T_{ex})_{\odot}$ is also applicable when the curve of growth for a star has been constructed from laboratory values of $g_i f_{ik}$.

Instead of comparing the atmosphere of a star with that of the Sun, one can with greater confidence compare the atmospheres of two stars of close spectral type, for example A and F supergiants, A and F dwarfs, etc.; this was first pointed out by O. A. Mel'nikov. In fact, it can be assumed in this case that both stars have the same atmospheric structure. This method has been applied in practice by T. M. Fofanova at Pulkovo to the study of a number of supergiants [42].

Returning again to the use of "solar" values of $\log_{10} X_f$, it must be pointed out that great caution has to be exercised. For, in determining $\log_{10} X_f$ from the curve of growth of the solar spectrum, it is implicitly assumed that all the "standard" lines here used correspond to the same absorbing (scattering) layer. At the same time, the presence of continuous absorption which varies with frequency may have the result that layers of varying "thickness" correspond to lines in different parts of the spectrum. This must be borne in mind. Hence, strictly speaking, it is best, in constructing curves of growth, to use laboratory values of $g_i f_{ik}$, or theoretical values if they are sufficiently reliable.

6. Curves of growth constructed from observation. Turbulent velocities in the atmospheres of stars. Let us now turn to the results obtained

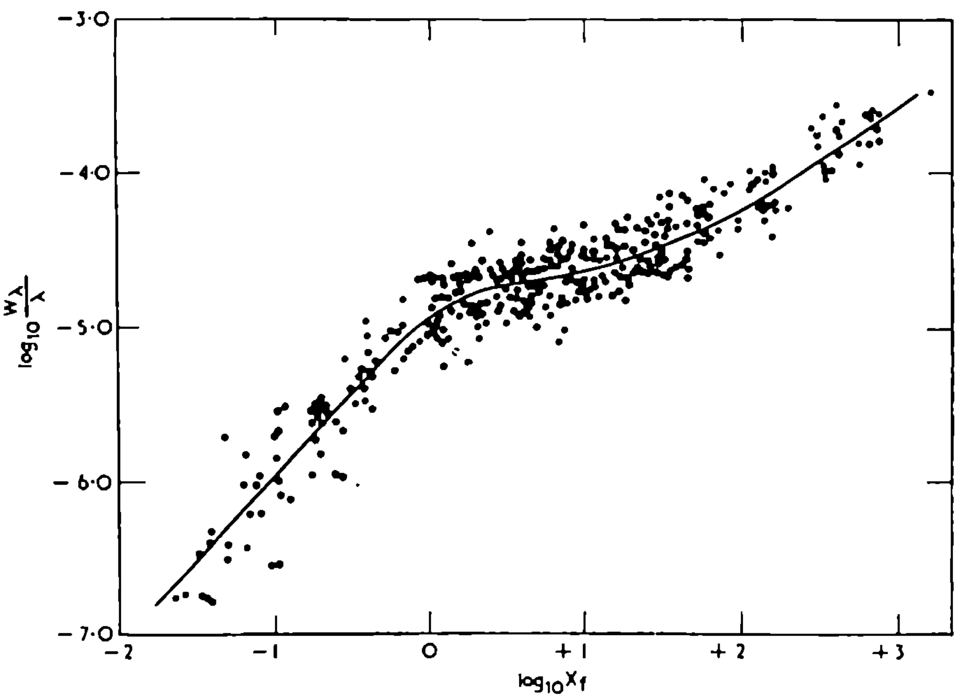

Fig. 32

by constructing curves of growth and to the interpretation of these results.

A very large amount of work has been devoted to the construction of curves of growth for the Sun and stars. In the Soviet Union such work is carried out at Pulkovo by O. A. MEL'NIKOV and his co-workers.

In Fig. 32 we give a curve of growth for the Sun, constructed at Pulkovo by L. A. MITROFANOVA from Fe I lines (the values of $g_i f_{ik}$ were determined in the laboratory from emission lines of Fe I). Fig. 33 shows a curve of growth for the supergiant α Cygni, constructed at Pulkovo by O. A. MEL'NIKOV from Ti II lines. Here again laboratory values of $g_i f_{ik}$ were used, as determined from emission lines. The points for each line are shown in Fig. 32, while in Fig. 33 the mean points (for several lines) are given, the diameter of the circle being proportional to the number of lines entering into the determination of the mean.

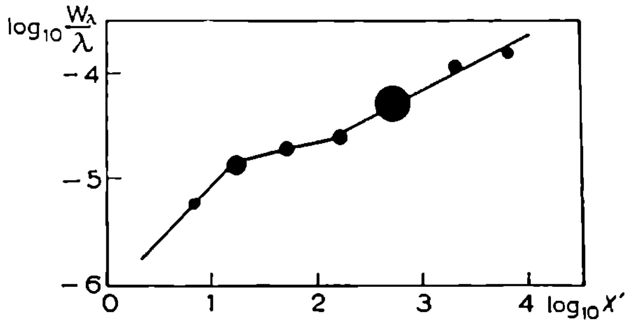

FIG. 33

As a result of this work, it has been possible to establish a number of interesting relations concerning the physical state of stellar atmospheres. The construction of curves of growth and their comparison with theoretical curves of growth has shown that in many cases, particularly in the atmospheres of supergiants, the value of v_0 is considerably larger than the mean kinetic (thermal) velocity corresponding to the mean temperature of the stellar atmospheres. Thus, for example, for supergiants of temperature comparable with that of the Sun, the turbulent velocity v_t, found from the curve of growth and the value of v_0 [formula (11.29) is used with $T \approx T_e$], reaches in some cases tens of kilometres per second, while the mean thermal velocity in this case is of the order of 1 km/sec. For the Sun itself, the turbulent velocities in the "reversing" layer are apparently not more than 1·5 km/sec. This result indicates that the conditions are more favourable to the appearance of turbulent currents in the very extended, relatively rarefied envelopes of supergiants, than in the dense non-extended atmospheres of dwarfs*.

* There are apparently exceptions (though as yet unexplained) to this rule. For example, the turbulent velocity in the supergiant α Carinae is approximately the same as in dwarfs. This shows that one must be cautious in making generalisations.

Furthermore, the existing data indicate that the value of v_t is larger for ionised than for neutral atoms. This is apparently because the mean level at which ionised atoms are found in the atmosphere of a supergiant is higher (since p_e is less) than the level for neutral atoms. Thus, for example, in the dwarf Procyon $v_t \approx 1$ km/sec for neutral atoms and $v_t = 3$ km/sec for ionised atoms. In the supergiant α Persei the corresponding numbers are $v_t \approx 3{\cdot}5$ and $6{\cdot}5$ km/sec. Finally, there are indications that in the atmospheres of supergiants v_t diminishes when the excitation potential ε_i increases. This, according to data on the solar chromosphere, corresponds to an increase of v_t when the level where the absorbing atoms are found becomes higher. As an example we may state that for α Persei the lines with $\varepsilon_i \approx 0{\cdot}0$ eV give $v_t = 7{\cdot}0$ km/sec, but those with $\varepsilon_i \approx 4{\cdot}5$ eV give $v_t \approx 3{\cdot}5$ km/sec. All these facts indicate the presence, in the atmospheres of supergiants, of different effective layers for different atoms. Ionised atoms lie considerably higher than neutral atoms.

The increase in the turbulent velocity with height in the atmosphere of the star is particularly marked in such binary systems as ζ Aurigae. The system ζ Aurigae consists of a giant of class K 5 with a very extended atmosphere, and a considerably smaller star of class B 8. Outside eclipse, the spectrum of the system is composite: in the red part the K spectrum is dominant, in the blue and violet part the B spectrum. When the B star passes behind the K star, numerous sharp absorption lines, formed by the absorption of light from the B star in the extended atmosphere of the red giant, appear in the blue and violet parts of the spectrum of the system, where the B spectrum has till then been dominant. By this means it becomes possible to study the nature of the atmosphere of the K star as a function of the height above its photosphere. When the B star emerges from behind the K star, the same events happen in the reverse order.

The construction of the curve of growth for various times after the emergence of the B star from behind the K star, i. e. for various heights, has led to the following results: for a height of $0{\cdot}8 \times 10^6$ km the turbulent velocity was equal to $6{\cdot}5$ km/sec, and for a height of $20{\cdot}6 \times 10^6$ km it was 13 km/sec. This directly confirms what was been said above.

In some cases the turbulent velocities are particularly high. Thus, for example, in ε Aurigae the turbulent velocity is apparently close to 20 km/sec, and the turbulent velocity in the atmosphere of 17 Leporis is even greater (67 km/sec).

It is clear that the presence of turbulent currents must lead to an increase in the line widths and to the obliteration of the lines. This effect is noticeable for large v_t even on merely glancing at a spectrogram of the star.

The presence of turbulent motions in the atmospheres of stars and the increase of turbulence with height indicate a complexity of the structure of the atmospheres of such stars. The atmospheres of the supergiants are marked by especial complexity. Here, besides the general turbulent motions, differential motions of various elements are observed. Thus, according to the investigations of G. A. SHAĬN and P. F. SHAĬN [146], the hydrogen lines in the spectra of the supergiants α Cygni and β Orionis have a positive displacement relative to the lines of ionised metals, amounting to $+ 1{\cdot}9$ km/sec for the former and $+ 4$ km/sec for the latter. In the spectrum of α Cygni it is probable that there is small displacement of the lines of ionised metals relative to the neutral metals, and so on. Here, besides the general turbulence, we apparently have certain currents of material with somewhat different velocities for different elements.

In concluding the subject of turbulent velocities, we must make the following remark. The study of absorption line contours in extended envelopes shows that they must differ from the case where the atmosphere is a thin layer (see Chapter 13). The appropriate analysis shows that, if this factor is not taken into account, the turbulent velocity obtained from the curve of growth is less than the actual turbulent velocity in the atmosphere of the star. For supergiants, whose atmospheres are extended, this effect may be considerable.

7. The excitation temperature. The damping constant. Let us now consider the determination of the excitation temperature. The methods of determining T_{ex} explained above are important not only from the viewpoint of finding T_{ex} itself, but also from that of testing the Boltzmann temperature distribution in the outer layers of stars. If the values of Y, and therefore, according to (12.5), (12.35), (12.37) and (12.38), the values of

$$\log_{10} \frac{N_i}{g_i} = Y - \log_{10} \frac{\sqrt{\pi}\, e^2}{m_e c} \frac{1}{v_0} = \log_{10} \frac{N_r}{u_r} - \frac{5040}{T_{ex}} \varepsilon_i\,, \quad (12.41)$$

are linear functions of ε_i, the application of Boltzmann's formula for the range of ε_i involved is justified. [The equality of the first and last members of formula (12.41) is Boltzmann's formula written in logarithmic form.]

The results of the corresponding investigations show that, for ε_i from 0 to about 5 eV, linearity holds in the majority of cases (of course, within the limits of observational error)*. As an example, we give a graph by O. A. MEL'NIKOV (Fig. 34), obtained from lines of neutral iron in the solar spectrum (u_0 is the partition function for neutral iron).

* For $\varepsilon_i > 5$ eV there are as yet no reliable results. See Chapter 15 concerning cases of deviation from the linearity mentioned.

However, when the temperature T_{ex} is determined from graphs of this kind, it is lower, in the majority of cases, than the effective temperature of the star. Thus, for example, T_e for the Sun $= 5710°$, while O. A. Mel'nikov obtained for the Sun $T_{ex} = 4580° \pm 100°$, using calibrated values of $g_i f_{ik}$ for Fe I. A. N. Demidova obtained $T_{ex} = 4700° \pm 450°$ using the theoretical $Ss/\sum s$ for Fe I, and S. E. Voĭnova obtained $T_{ex} = 4780° \pm 15°$ from the mean curve of growth for Fe I. Similar anomalously low values of T_{ex} for the Sun have been obtained from other data, although in some cases, such as, for instance, V I, a value $T_{ex} \approx 5100°$ is obtained, which is closer to $T_e \approx 5700°$. Moreover, the aggregate of all values of T_{ex} for the Sun shows a very large dispersion, as regards T_{ex}, among different elements. It is possible that here T_{ex} is actually different for different elements.

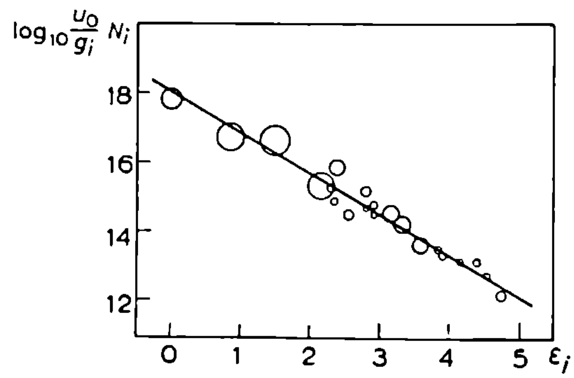

Fig. 34

Similar results ($T_{ex} < T_e$) are usually obtained for stars, and the difference between T_e and T_{ex} may sometimes amount to several thousand degrees*. Moreover, the difference $T_e - T_{ex}$ is apparently less for giants than for dwarfs. Thus, for example, for Sirius and γ Geminorum (stars of the main sequence), where $T_e \approx 10{,}000°$, L. H. Aller finds $T_{ex} \approx 6000°$.

We may remark that, for stars, T_{ex} can be determined either from the linear dependence of Y on ε_i, or by means of a comparison of line intensities in the star's spectrum with the corresponding lines in the Sun's spectrum [see formula (12.40)].

In conclusion, let us consider briefly the interpretation of the inequality $T_{ex} < T_e$. The position here is not yet altogether clear, but two possibilities must be mentioned: (1) an attempt might be made to explain the anomalously low T_{ex} by using the fact that, in the frequencies of absorption lines, where the processes of excitation of atoms take place, the density of radiation is lessened by the presence in these frequencies of the absorption lines themselves. However (as we shall see in Chapter 15), this explanation must apparently be rejected from the theoretical point of view. (2) A detailed analysis of curves of growth shows that an increase of v_t, i. e. the turbulent velocity, with the excitation potential ε_i has the same effect on the curve of growth, in a finite interval of $\log_{10} X_0$, as a decrease of T_{ex}. In other words, one

* According to O. A. Mel'nikov [82], the same result is obtained for cepheids, which are non-steady stars.

of these factors can compensate the other. Let us now assume that v_t increases when ε_i decreases. Then, on making the erroneous assumption that the scattering of points round the curve of growth is due only to errors in W_λ and f_{ik}, i. e. assuming v_t = constant, we obtain too low a value for T_{ex}. It is possible that the difference between T_{ex} and T_e may be partly due to this circumstance, since, as we have seen, v_t in fact increases when ε_i decreases, i. e. with increasing height in the stellar atmosphere. This explanation (proposed by O. STRUVE and A. UNSÖLD) can be tested only on multiplets whose intensities cover the entire curve of growth.

Let us now discuss the values of Γ_{ik}/ν_0 obtained from the curve of growth. In the majority of cases the value of Γ_{ik}/ν_0, which, as we have said, is taken as a mean for the whole curve of growth, is considerably greater than the value which corresponds only to radiation damping, which is on the average about $1{\cdot}5 \times 10^{-7}$ (for $\lambda = 5000$ Å). Thus, for instance, for the Sun, the mean value of Γ_{ik}/ν_0 from a large number of investigations is found to be some 5 to 10 times greater than the value mentioned, $1{\cdot}5 \times 10^{-7}$. The same thing is found for the majority of dwarf stars. These results can be explained only by the fact that the constant δ_{ik} in formula (11.25), and therefore Γ_{ik} also, is mainly determined not by radiation damping, but *by collision damping*.

We shall confirm this important result in the next chapter. Here we can test the result (though only qualitatively), starting from the following considerations.

If the value of Γ_{ik}/ν_0 is in fact determined by collision damping, then from formulae (11.45), (11.47) and (11.49) this quantity should increase and decrease with the density of the material. In other words, it is to be expected that Γ_{ik}/ν_0 is smaller in the atmospheres of giants than in those of dwarfs. For ionised atoms it should be, on the average, less than for neutral atoms, since the former lie, on the average, higher than the latter. These suppositions are partly confirmed by observation, although in some cases there are exceptions to the expected relations. Thus, it follows from the existing data that for *neutral elements* in supergiants the value of Γ_{ik}/ν_0 is fairly close to the value for dwarfs, i. e. 10 times greater than the radiation damping constant, but for *ionised elements* Γ_{ik}/ν_0 is close to the value $1{\cdot}5 \times 10^{-7}$.

In general, the material at present existing is meagre, and further investigations are necessary here before final conclusions can be reached.

8. **Critical remarks.** The rational use of curves of growth can lead to a number of important results. However, it must be kept in mind that the theory developed above is very simplified and schematic. For this reason a number of critical remarks must be made. Firstly, caution must be exercised in the use of observational data. Thus, for instance,

T. M. Fofanova (at Pulkovo) has shown [42] that, when spectrographs of small dispersion are used, the velocity v_0 found from the curve of growth increases (*spuriously*) when the dispersion of the spectrograph used decreases. Hence the spectrograms used to construct curves of growth must have a fairly large linear dispersion (better than 5 to 10 Å/mm). Secondly, the theory of curves of growth developed above is based on a large number of simplifying assumptions, of which the main ones are as follows:

(1) The model of the "reversing layer" which we have used, and the formula (12.9) corresponding to it, are very much simplified. In particular, this formula presupposes the same "thickness" of the reversing layer for all λ, which is equivalent to excluding any dependence of $\varkappa_\nu$ on ν. In this connection, a number of calculations have recently been made which were based on the more exact formula (10.20) and its various modifications.

The expression for η_ν, by virtue of formula (11.39) and the fact that $\sigma_\nu/\sigma_{\nu_0} = s_\nu/s_{\nu_0}$, can be rewritten as

$$\eta_\nu = \sigma_\nu/\varkappa_\nu = (\sigma_{\nu_0}/\varkappa_\nu)\,\phi(a,p) = \eta_{\nu_0}\,\phi(a,p)\,, \tag{12.42}$$

where

$$\eta_{\nu_0} = \sigma_{\nu_0}/\varkappa_\nu\,. \tag{12.43}$$

The argument X_0 of the elementary theory of the curve of growth is thus replaced by η_{ν_0}. We again obtain a family of curves of growth for various values of the constant a. In constructing such a family of curves, a fixed value must be taken for the ratio b_ν/a_ν, and for the quantity ε_ν which appears through (10.9) in formula (10.20). As we shall see in Chapter 14, in many cases we can take $\varepsilon_\nu = 0$ for the calculation of equivalent widths.

The results of the corresponding calculations show that the general character of the curves of growth constructed by means of formula (10.20) is the same as in Figs. 30 and 31, although there are certain differences of detail. However, a comparison of the curves of growth constructed by T. M. Fofanova [42]* for various models shows that these differences cannot be investigated with reasonable certainty with the present dispersion of the separate points round the curve of growth. Nevertheless, the problem of constructing a physically correct theoretical curve of growth is one of the main tasks of theoretical astrophysics.

* In this work the influence of the change in the coefficient of continuous absorption with wavelength on the parameters found from the curve of growth is studied. This influence appears to be comparatively small.

(2) In the elementary theory which we have considered, some mean value of Γ_{ik}/ν_0 was taken for all lines. This assumption, however, is a crude one both in the case of radiation damping and in that of collision damping. Moreover, in the theory explained above it is supposed that Γ_{ik}/ν_0 does not depend on the depth in the stellar atmosphere. In the case where Γ_{ik}/ν_0 is determined by collision damping, this is obviously incorrect, since the value of Γ_{ik} depends on the number of particles at the point considered.

(3) Furthermore, in constructing the composite curve of growth from several elements, a mean value is taken for v_0; this is usually justified on the grounds that the turbulent velocity is about the same for all atoms, which is also an approximation.

(4) Finally, it is assumed in the theory explained above that the excitation temperature T_{ex} is constant throughout the star's atmosphere. This, as we shall see in Chapter 15, is likewise a fairly crude approximation.

Hence it is not surprising that the agreement between the theoretical curves of growth and those derived from observation is very often unsatisfactory. The discordances in the direction of the X_0 axis sometimes reach 0·15 to 0·20 in the logarithm of X_0, or, what is the same thing, in the logarithm of $N_i f_{ik}$. The use of formula (10.20) instead of (12.9) does not appreciably change the situation.

Chapter 13. The interpretation of observed contours of absorption lines in stellar and solar spectra

1. **The calculation of the ratio of the coefficients of selective and continuous absorption.** In the present chapter we shall consider the subject of absorption line contours in the spectra of stars of various classes. We shall first consider lines of metals, and then those of hydrogen and helium. We shall also discuss the effect of the star's rotation on the line contours in its spectrum.

The theoretical calculations relating to resonance lines are at present the most trustworthy*. For lines which begin from excited levels (subordinate lines) the calculations are more indefinite (see Chapter 14). The chief difficulty arises from the fact that the processes of absorption and subsequent re-emission are completely non-coherent for subordinate lines.

* This remark does not refer to the theory of contours distorted by the rotation of the stars.

The construction of a line contour amounts in the first place to the calculation of the quantity η_ν:

$$\eta_\nu = \sigma_\nu / \varkappa_\nu \,, \tag{13.1}$$

which appears in the equation of transfer (10.2). The quantity ε_ν is less important, and usually plays a small part except in the central frequencies of the line. If all three quantities η_ν, ε_ν and B_ν in equation (10.2) are expressed in terms of the fundamental physical parameters T, p, p_e, ϱ, etc., and the latter are related to the optical depth τ_ν for the frequency of the line (as in Chapter 7, using the theory of photospheres), the problem reduces simply to the solution of equation (10.2). Bearing these remarks in mind, let us consider an actual example: the absorption lines in the Sun's spectrum. (Hotter stars will be considered later.)

We shall suppose throughout that the density of matter ϱ to which the coefficients σ_ν and $\varkappa_\nu$ in equation (9.29), and therefore in (10.2), relate, is the *total* density of matter*, taking account of *all* the atoms in 1 cm^3. We shall determine the value of σ_ν, starting from the general formula (5.23).

Let s_ν be the absorption coefficient for an atom in the ith excitation state and rth ionisation state. The number of such atoms in 1 cm^3 is $n_{r,\,i}$. In accordance with (5.23), we can write for the selective attenuation of the intensity by the atoms in question

$$\mathrm{d}I_\nu = -\,I_\nu\, s_\nu\, n_{r,\,i}\, \mathrm{d}h = -\,I_\nu\, s_\nu\, \frac{n_{r,\,i}}{\varrho}\, \varrho\, \mathrm{d}h \,. \tag{13.2}$$

Consequently, the selective absorption coefficient, referred to 1 gram of total mass, is

$$\sigma_\nu = s_\nu\, n_{r,\,i}/\varrho \,. \tag{13.3}$$

For subsequent work it is useful to have the number of atoms in the ith and rth states referrred not to 1 cm^3 but to 1 *gram* of matter (taking account of atoms of all elements); this number, which we denote by $n_{r,i}^{(g)}$, is

$$n_{r,\,i}^{(g)} = n_{r,\,i}/\varrho \,. \tag{13.4}$$

Consequently, the general relation

$$\sigma_\nu = s_\nu\, n_{r,\,i}^{(g)} \tag{13.5}$$

* We might also refer σ_ν and $\varkappa_\nu$ to the density determined by hydrogen atoms, but the method used is more convenient.

holds. Next, let the *total* number of atoms of the element considered, in all states of ionisation and excitation, referred to 1 g of matter, be $n^{(g)}$. We can relate $n_{r,i}{}^{(g)}$ and $n^{(g)}$ as follows:

$$n_{r,i}{}^{(g)} = Z_{r,i}\,(p_e, T)\, n^{(g)}\ , \tag{13.6}$$

where the function $Z_{r,i}$ can easily be calculated from Saha's and Boltzmann's formulae (5.11) and (5.54). [For an example of the calculation see Chapter 16, formulae (16.1)—(16.11).]

In turn, the quantity $n^{(g)}$ can be determined as follows. Let n be the *total* number of atoms of the element considered in 1 cm³. Then it is evident that

$$n^{(g)} = n/\varrho\ . \tag{13.7}$$

We can write the density ϱ as a sum:

$$\varrho = \sum_s n_s\, m_s\ , \tag{13.8}$$

where m_s is the mass of an atom of the corresponding element, and the summation is to be extended over *all* elements. The expression (13.8) can also be written

$$\varrho = n_{\mathrm{H}}\, m_{\mathrm{H}} \left\{ 1 + \frac{n_{\mathrm{He}}\, m_{\mathrm{He}}}{n_{\mathrm{H}} m_{\mathrm{H}}} + \frac{n_{\mathrm{Li}}\, m_{\mathrm{Li}}}{n_{\mathrm{H}}\, m_{\mathrm{H}}} + \ldots \right\}, \tag{13.9}$$

where we have added the symbols of the corresponding elements as suffixes to the numbers n and m. As may easily be calculated from Table 1, only the first two terms in the braces in (13.9), which take account of *hydrogen* and *helium*, are important. If, using Table 1, we take $n_{\mathrm{He}}/n_{\mathrm{H}} = 0{\cdot}2$, the quantity in the braces is equal to 1·8. If we take more recent estimates (see Chapter 15), which give $a_s \approx 0{\cdot}05$ for helium in B and O stars, the quantity in the braces is equal to 1·2. However, we shall, for the sake of generality, denote this quantity by b. Using (13.5), (13.6), (13.7) and (13.9), we obtain for σ_ν

$$\sigma_\nu = s_\nu\, Z_{r,i}\,(p_e, T)\, \frac{1}{b\, m_{\mathrm{H}}}\, \frac{n}{n_{\mathrm{H}}}\ . \tag{13.10}$$

Thus, to calculate σ_ν, it is necessary to know the content of the given element relative to hydrogen. We denote this content by a_s, in accordance with the notation of Chapter 5. The quantity n_1 in formula (5.1) is in our case the number of *hydrogen atoms* in 1 cm³ (both neutral and ionised).

The content of atoms of some metal in some volume is often given relative to the total number of atoms of all the metals in this volume, and is denoted by α. The content of hydrogen atoms relative to all

metal atoms is denoted by A; the value of A in stellar atmospheres is close to 10,000. Thus, if n_{M} is the number of all metal atoms in 1 cm^3, formula (13.10) can be rewritten

$$\begin{aligned} \sigma_\nu &= s_\nu Z_{r,i}(p_e, T) \frac{a_s}{b\, m_{\mathrm{H}}} \\ &= s_\nu Z_{r,i}(p_e, T) \frac{1}{b\, m_{\mathrm{H}}} \frac{n}{n_{\mathrm{M}}} \frac{n_{\mathrm{M}}}{n_{\mathrm{H}}} \qquad (13.11) \\ &= s_\nu Z_{r,i}(p_e, T) \frac{1}{b\, m_{\mathrm{H}}} \frac{\alpha}{A} . \end{aligned}$$

The methods of determining the values of a_s, α and A will be considered in Chapter 15. At present we can take them as given, say from Table 1.

We now turn to the coefficient of *continuous* absorption $\varkappa_\nu$ which appears in formula (13.1). In the solar atmosphere negative hydrogen ions play the chief part as far as a mean optical depth $\bar{\tau}$ of the order of unity. On penetrating further, the absorption by neutral hydrogen atoms becomes more and more important. Thus we can write

$$\varkappa_\nu = \varkappa_\nu(\mathrm{H}^-) + \varkappa_\nu(\mathrm{H}) . \qquad (13.12)$$

Let us consider $\varkappa_\nu(\mathrm{H}^-)$. The absorption coefficient for H^- is usually referred to one *neutral* hydrogen atom and unit electron pressure (1 bar). This coefficient, which is shown graphically in Fig. 15(b), is denoted by $k_\nu(\mathrm{H}^-)$. Using again formula (5.23), as we did in finding σ_ν from (13.2), we have

$$\varkappa_\nu(\mathrm{H}^-) = \frac{(n_0)_{\mathrm{H}}\, k_\nu(\mathrm{H}^-)}{\varrho}\, p_e , \qquad (13.13)$$

where $(n_0)_{\mathrm{H}}$ is the number of neutral hydrogen atoms in 1 cm^3. Introducing the expression (13.9) into (13.13), we find

$$\varkappa_\nu(\mathrm{H}^-) = \frac{(n_0)_{\mathrm{H}}}{n_{\mathrm{H}}} \frac{k_\nu(\mathrm{H}^-)}{b\, m_{\mathrm{H}}}\, p_e . \qquad (13.14)$$

In the solar photosphere as far as $\bar{\tau} = 5$ or 7, the hydrogen is practically all neutral, so that $(n_0)_{\mathrm{H}}/n_{\mathrm{H}} \approx 1$.

We now turn to the coefficient $\varkappa_\nu(\mathrm{H})$. The continuous absorption coefficient for hydrogen, given by formula (5.68), is referred to one hydrogen atom (Z is unity for hydrogen). We denote it by $k_\nu(\mathrm{H})$. Proceeding as in the derivation of (13.13), we obtain for $\varkappa_\nu(\mathrm{H})$

$$\varkappa_\nu(\mathrm{H}) = k_\nu(\mathrm{H})\, n_{\mathrm{H}}/\varrho . \qquad (13.15)$$

Again using (13.9), we find

$$\varkappa_\nu(\mathrm{H}) = k_\nu(\mathrm{H})/b\, m_{\mathrm{H}} . \qquad (13.16)$$

Thus, taking into account that $(n_0)_{\rm H}/n_{\rm H} \approx 1$, we obtain for $\varkappa_\nu$

$$\varkappa_\nu = \varkappa_\nu(\mathrm{H}^-) + \varkappa_\nu(\mathrm{H}) = \frac{1}{b\, m_{\rm H}}\,[k_\nu(\mathrm{H}^-)\, p_e + k_\nu(\mathrm{H})]\,. \qquad (13.17)$$

Finally, from (13.11) and (13.17) we find for η_ν

$$\eta_\nu = \frac{\sigma_\nu}{\varkappa_\nu} = \frac{s_\nu\, a_s\, Z_{r,i}(p_e, T)}{k_\nu(\mathrm{H}^-)\, p_e + k_\nu(\mathrm{H})}\,. \qquad (13.18)$$

If the value of s_ν is given, then η_ν is a known function of the fundamental physical parameters of the atmosphere, T, p, p_e and ϱ. In turn, the variation of these with depth should be known from the theory of stellar photospheres. The variable τ_ν which appears in the equation of transfer (10.2) is sometimes replaced by the *mean* optical thickness $\bar{\tau}$, which we defined in Chapter 6; the following ratio is then introduced:

$$n_\nu = \varkappa_\nu/\bar{\varkappa}\,, \qquad (13.19)$$

where $\bar{\varkappa}$ is the absorption coefficient $\varkappa_\nu$ averaged in some manner over the whole spectrum (see Chapter 6). In this case $d\tau_\nu$ in equation (10.2) takes the form

$$d\tau_\nu = \varkappa_\nu\, \varrho\, dh = n_\nu\, \bar{\varkappa}\, \varrho\, dh = n_\nu\, d\tau\,. \qquad (13.20)$$

This method is convenient, because in the solar photosphere n_ν depends only slightly on depth. Hence we can put approximately $n_\nu \approx$ constant. In this case

$$\tau_\nu = n_\nu\, \tau\,. \qquad (13.21)$$

Let us now consider the quantity ε_ν in equation (10.2). The value of ε_ν in a large number of cases, and in particular for resonance lines, satisfies the inequality $\varepsilon_\nu \ll 1$. In such cases, as is easily seen, for instance, from formulae (10.20) and (10.9), the effect of the value of ε_ν on the intensity can be noticeable only for the *innermost* parts of the lines, i. e. for the core. For the wings of lines, where η_ν is less than or comparable with unity, the part played by ε_ν is negligibly small. On the other hand, the general problem of *central* residual intensities has not yet been solved; these intensities have observed values which are higher than those given by theory. In the present chapter we shall concern ourselves only with the *outer* parts of absorption line contours, and we shall not treat the problem of the central parts of lines. For this reason we shall at present assume that $\varepsilon_\nu = 0$.

2. **The application of the theory to solar absorption lines. The part played by pressure effects.** In order to construct the contour of some absorption line, using (13.18) and the assumption that $\varepsilon_\nu = 0$, we must

fix the value of δ_{ik} [determined by formula (11.25)] in the expression (11.39) or (11.24) for s_ν.

As an actual example, let us consider the D lines of neutral sodium. Since the D lines are *resonance* lines, we can assume that in (11.25) $\tau_i = \infty$ and consequently $\gamma_i = \gamma_1 = 0$. Also, for the D lines under solar conditions, the last three terms on the right-hand side of (11.20) are negligibly small compared with the first term, i. e. A_{21}. Hence we can write for δ_{ik}

$$\delta_{12} = \frac{\Gamma_{12}}{4\pi} = \frac{1}{4\pi}(\gamma_2 + \gamma_c) \approx \frac{1}{4\pi}(A_{21} + \gamma_c)\,. \tag{13.22}$$

Let us now calculate γ_c. As is shown by calculations of $\Delta\lambda_g$, the contour of practically the whole line is determined only *by collision damping.* In order to decide which of the two cases ($k = 4$ or $k = 6$) is to be applied to the D lines of the solar spectrum, it is necessary to compare the corresponding values of γ_c, given by formulae (11.47) and (11.49). It can be predicted that the case where $k = 6$ is the more important. The lines which are particularly sensitive to the quadratic Stark effect are those connected with highly excited, weakly bound terms (namely the D, F, etc. terms); the D lines, on the other hand, are resonance lines with S—P transitions. Bearing this in mind, we shall start with γ_c for $k = 6$.

For the calculations we shall take a level in the solar atmosphere corresponding to an optical depth $\tau_\lambda = 0{\cdot}48$ for $\lambda = 5010$ Å. For this level, the temperature T is 5895°, calculated from the theory of photospheres (see Table 13, Chapter 17) with $T_e = 5710°$; also, for this level $p \approx 1{\cdot}32 \times 10^5$ bars, while $p_e \approx 19$ bars.

We now apply formula (11.49). For $\bar{v}$ we obtain for sodium and hydrogen $1{\cdot}1 \times 10^6$ cm/sec. Also, as we have already said, $C = 2{\cdot}65 \times 10^{-32}$ for the D lines. For p, we can write with sufficient accuracy:

$$p = p_{\rm H} + p_{\rm He} = (n_{\rm H} + n_{\rm He})\, kT \approx 1{\cdot}2\, n_{\rm H}\, kT\,, \tag{13.23}$$

where we have taken the value 0·2 for the ratio $n_{\rm He}/n_{\rm H}$, in accordance with Table 1. For the values of T and p which we have taken, $n_{\rm H} = 1{\cdot}3 \times 10^{17}$. Introducing these values into (11.49), we find that $\gamma_c = 2{\cdot}2 \times 10^9$. On the other hand, for the D lines $\gamma_2 \approx A_{21}$, which is of the order of 7×10^7. Consequently, for the D lines at the depth in the solar atmosphere considered, γ_c is approximately 30 times greater than the radiation damping constant. When τ_λ decreases, γ_c decreases fairly slowly. Thus, for $\tau_\lambda = 0{\cdot}15$ the value of γ_c is diminished by a factor of about 0·6, but even so it is about 20 times greater than $\gamma_2 = A_{21}$. It is true that the value of C is approximate, as we have said. However, despite the possible errors, it cannot be doubted that γ_c exceeds γ_2 by at least an order of magnitude.

The broadening caused by electrons is in this case considerably less. Laboratory measurements have shown that the D lines are displaced by $\Delta\lambda = 0{\cdot}025$ Å at 160 kV/cm. Using formula (11.44) to determine C, and formula (11.47) with $n = n_e$ for γ_c, it is easily found that, for the value of $p_e = 19$ bars given above, the value of γ_c is of the order of $5{\cdot}4 \times 10^7$, i. e. even less than $\gamma_2 = A_{21}$. Thus the damping by collisions of sodium atoms with hydrogen atoms is the chief factor which broadens the D lines in the solar spectrum. In fact, this result apparently holds for all resonance lines in the solar spectrum.

Having established the chief factors which lead to the broadening of the lines in which we are interested (as we have just done for the example of the D lines of Na I), we can determine δ_{ik} by formula (11.25), and thus [by formula (11.38)] the value of a for every depth in the solar atmosphere. The ratio η_ν can be determined by means of formulae (11.39) and (13.18):

$$\eta_\nu = \frac{a_g Z_{r,i}(p_e, T)}{k_\nu(\mathrm{H}^-)\, p_e + k_\nu(\mathrm{H})}\, \phi\,(a, p)\, s_{\nu_0}, \tag{13.24}$$

where s_{ν_0} is given by formula (11.40).

We must first study the dependence of η_ν, for the given line, on depth in the solar atmosphere. The fundamental parameters T, p, p_e and ϱ are given by the theory of the photosphere. If the change of η_ν with depth is comparatively small ($\eta_\nu \approx$ constant), and the dependence of B_ν on τ_ν can be represented with sufficient accuracy in the form of a linear function [formula (8.10)], then we can use formula (10.20) for the line contour*; we have already said that for the wings of resonance lines we can put $\varepsilon_\nu = 0$. If η_ν changes markedly with depth, it is best to use the method of numerical integration of the equation of transfer or the method of iteration (see Chapter 10).

For the outer parts of the line wings and for faint lines, and in general for cases where the line depth R_ν does not exceed 15 %, we can use formula (10.43). We recall that in this formula the functions $G_1(\tau_\nu)$ and $G_2(\tau_\nu)$ can be obtained directly from the theory of photospheres, or, for the Sun, from the law of darkening of the disc to the limb. In our case ($\varepsilon_\nu = 0$) the true selective absorption coefficient $\varkappa_\nu^0$ is equal to zero by (9.31), and $\sigma_\nu^0 = \sigma_\nu$. Consequently, the first term in the last member of formula (10.43) is equal to zero, and the second contains in the integrand the ratio $\sigma_\nu^0/\varkappa_\nu = \sigma_\nu/\varkappa_\nu = \eta_\nu$, which we have already estimated [formula (13.24)]. Thus, to find R_ν, it suffices to estimate numerically the second integral in the last member of formula (10.43).

* For the case where η_ν varies only slightly with depth, and the deviations from the linear expansion (8.10) are small, the corresponding formulae have been given by several authors. For these formulae the reader is referred to the literature (see, for instance, M. Tuberg [166]).

The construction of line contours from the formulae mentioned has been carried out here for the whole disc of the Sun or star, using the *flux* of radiation. The same can also be done for the separate points of the solar disc, using the *intensity* of radiation. In this case the equation of transfer (10.2) must be solved for a given angle θ, and not for the flux. This can be done, in particular, for all the three cases which we considered in Chapter 10:

(1) The "standard" case with η_ν = constant, ε_ν = constant and $B_\nu = a_\nu + b_\nu \tau_\nu$.
(2) Numerical integration of the equation of transfer.
(3) The method which we considered in Chapter 10 for small R_ν [168, p. 384].

In these cases also the ratio η_ν is the starting-point; we have already considered the general methods of obtaining it.

Having constructed the contour of some line for various θ, we can then study the change of this contour with distance from the centre of the solar disc, and also compare the results with observation. In Fig. 35 we give such a comparison for the D_1 line of the solar spectrum,

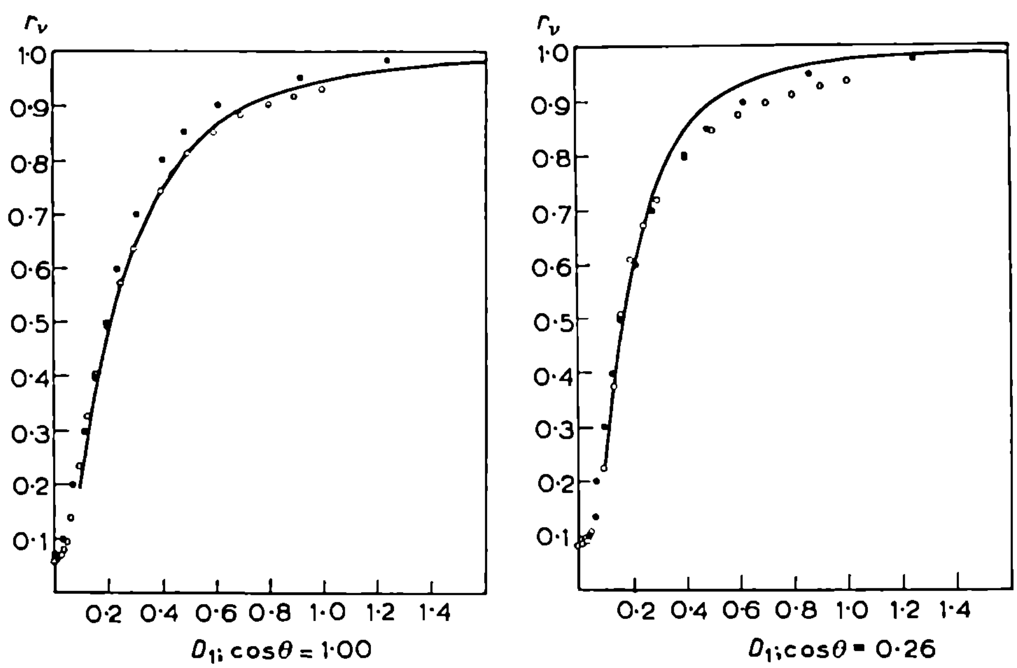

Fig. 35

due to D. L. Harris [58]. The abscissae in Fig. 35 are in ångströms. The value of η_ν is here calculated in the manner described above, and the contour is constructed for the case where the change of η_ν in the atmosphere is small in comparison with η_ν itself, and B_ν is a linear function of τ_ν.

In the case considered, these conditions are fulfilled with sufficient accuracy. The coefficients a_ν and b_ν in the expansion (8.10) are taken from the observed law of darkening. The contours have been constructed for two points on the solar disc; for the centre, and for $R = 0{\cdot}97\ R_\odot$. For the relative content of sodium the value $a_s \approx 3\times 10^{-6}$ was taken; this is close to the a_s given in Table 1 ($\approx 2\times 10^{-6}$). Fig. 35 shows that the theoretical contour, in the calculation of which collision damping was taken into account, is largely in satisfactory agreement with the contour obtained from observation, although there are also definite discrepancies between them, both in the wings and in the centre. In particular, the theory in question with $\varepsilon_\nu = 0$ predicts $r_{\nu_0} \approx 0$ in the centre of the line, whereas the observations give $r_{\nu_0} \approx 0{\cdot}06$. The cause of these discrepancies is not yet clear. It is quite possible that the decisive part here is played by processes of non-coherent scattering, which we shall discuss in the next chapter.

The comparison of theory and observation leads to much the same results for other resonance lines in the solar spectrum. Thus, despite the small discrepancies which still remain, the theory of absorption lines, taking pressure effects into account, is quantitatively in fairly satisfactory agreement with observational results. If we were to neglect pressure effects and assume that the value of δ_{ik} is determined only by radiation damping, the theory would give much narrower lines than those observed. This important result is in complete agreement with those regarding Γ_{ik} obtained from the analysis of curves of growth.

In addition to the results enumerated which relate to pressure effects, one more fact must be mentioned which indicates the great importance of these effects. For instance, it is found that the values of W_λ/λ for all the measured lines of the diffuse series $3^1\mathrm{P}^0 - n^1\mathrm{D}$ of neutral magnesium in the Sun's spectrum are practically the same. They all lie on the same part of the curve of growth, corresponding to radiation damping. In this case, by formula (12.28),

$$W_\lambda/\lambda \sim \lambda \sqrt{(f_{ik}\,\delta_{ik})}\,. \tag{13.25}$$

When the number of the line in the series increases, the value of f_{ik} decreases. If, further, δ_{ik} were determined only by radiation damping, then it would diminish when k increased, by (11.25) and (11.20), since A_{ki} decreases when k increases. Finally, the value of λ also decreases when k increases. Thus, in the case considered, the values of W_λ/λ for the lines of the given series would have to diminish when the number of the line in the series increased. Hence the observed constancy of these values shows that γ_c increases with the serial number of the line, and this increase compensates the decrease of f_{ik}, λ and A_{ki}. The increase

of γ_c with the serial number of the line is easily explained. The greater k, the larger the orbit of the electron in the atom and the larger the effective cross-section for broadening collisions. Calculations show that the damping constant γ_c increases by a factor of about 30 or 40 when the serial number k goes from 3 to 9.

Thus, in the majority of cases, pressure effects play a much more important part in the Sun's atmosphere than radiation damping, so that very often $\gamma_c \gg \gamma_i + \gamma_k$. However, the problem has not yet been finally resolved. In particular, the constants C in formulae (11.45), (11.47) and (11.49) are known for only a small number of lines, and even then the numerical values are only approximate. It is also necessary to exercise great caution in deciding which case of broadening plays the chief part, that with $k = 4$ or that with $k = 6$. Sometimes both cases must be taken into account. Thus, for instance, in the $3^1\mathrm{P}^0 - n^1\mathrm{D}$ series of Mg I, the main effect in the first few terms is due to neutral hydrogen atoms ($k = 6$), but for the higher terms the broadening of the lines by electrons (and to a lesser degree by ions — see Chapter 11) begins to play the chief part, i. e. the case $k = 4$ becomes important. Thus, for some lines of the series, both effects are of about the same importance.

3. Lines of metals in the spectra of stars. Line broadening by turbulence. Let us now discuss briefly the subject of the contours of metallic lines in the spectra of stars of other classes. The methods used here are in general the same as in the case of the Sun. We must calculate the structure of the photosphere of the star considered and obtain the appropriate expressions for η_ν and for ε_ν (see Chapter 14). The problem then reduces to the integration of the equation of transfer by some method. However, in the calculation of line contours in the spectra of stars, the subsidiary problem very frequently arises of determining the acceleration g due to gravity at the surface of the star. Without knowing this quantity, we cannot calculate the structure of the photosphere. If the mass and radius of the star in question are unknown, the value of g has to be found from absorption lines. One of the most reliable methods here is the comparison of the theoretical contours of the Balmer series with those in the spectrum of the star for which g is sought. Other lines can also be used, besides those of the Balmer series. The effective temperature T_e of the star must be determined by the methods of Chapter 16.

Sufficiently reliable theoretical contours of absorption lines of metals in the spectra of stars differing markedly from the Sun are not yet available. Hence it is impossible to speak of any conclusions from the comparison of theory and observation. The pressure effects are the most interesting in the construction of contours.

For stars considerably cooler than the Sun, the construction of theoretical contours is complicated by the presence of molecular bands. For fairly hot stars, where hydrogen is ionised to a considerable extent, the contour of a line is determined by the collisions of atoms with ions and electrons and particularly with atoms of neutral helium. At still higher temperatures, helium also becomes completely ionised.

Numerous observations have shown that the lines of ionised metals in the spectra of supergiants are, as a rule, deeper and broader than the same lines in the spectra of fainter stars of the same class. (The reverse is observed for hydrogen and helium lines.) This fact has not yet been explained; the investigations are hampered by the lack of sufficiently accurate contours of metallic lines. Because of the smaller acceleration due to gravity in the atmospheres of supergiants, compared with the atmospheres of stars of smaller absolute luminosity, the densities in the former will be lower than in the latter. This must affect the coefficients of both continuous and selective absorption.

In the spectra of some supergiants the lines of metals are so much broadened that they immediately suggest the presence, in the atmospheres of these stars, of turbulent velocities, of which we have already spoken in the last chapter. In some cases this supposition is confirmed by the fact that the cores of the absorption lines can be represented by a Doppler distribution (11.34), but with $\Delta\nu_D$ corresponding to velocities of the order of tens of kilometres per second. The curves of growth for such stars usually indicate a v_t of the order of a few kilometres per second. Thus, for example, on analysing the line contours in the spectrum of δ Canis Majoris, O. Struve found a turbulent velocity of the material of about 30 km/sec, while the curve of growth for this star gives a turbulent velocity of the order of 5 km/sec.

The most probable explanation of this fact is as follows. On the basis of present-day ideas on the nature of turbulence, we have reason to suppose that in stellar atmospheres turbulence is characterised by the presence of separate cells of various dimensions, moving in various directions with various velocities. In order to make the sequel more easily intelligible, let us consider two limiting cases. In the first of these, we suppose that the dimensions of the turbulence cells are *comparable* with or *greater* than the thickness of the "reversing layer". In the second, we suppose that the cell dimensions are *considerably less* than this thickness.

It is quite evident that in the first case every cell in the reversing layer of the star absorbs photospheric radiation *independently* of the presence of other turbulence cells. Hence the existence of turbulent motions does not here increase the amount of photospheric energy absorbed in the line, i. e. the equivalent width of the line. The line is

only broadened correspondingly, its contour being determined by the velocity distribution of the cells (as well as by the number of absorbing atoms). Thus, in the case considered, the presence of turbulence, which broadens the absorption lines, cannot change the curve of growth constructed for $v_t = 0$. This case is very similar to that of a rotating star, where the values of W_λ are again unaltered.

In the second case, the situation is very similar to that which exists in the presence of ordinary thermal motions of Doppler character. In fact, the separate cells here lie one upon another. Hence a ray passing through the reversing layer outwards (from the photosphere) is weakened by absorption in many cells having approximately the same component velocities in the direction of the ray. Consequently, in the case considered, the presence of turbulence increases the equivalent width, and hence alters the curve of growth. This case was, in essentials, considered in Chapter 12.

The considerations just given account completely for the nature of the spectra of such stars as δ Canis Majoris. In fact, the large turbulent velocity obtained from the contour and the considerably smaller one obtained from the curve of growth indicate that the thickness of the reversing layer in these stars is *less* than the most frequent dimensions of the turbulence cells* in the atmospheres of the stars.

The fact that in the atmospheres of such stars as ε Aurigae and 17 Leporis the value of v_t, found directly from the curves of growth, is large, is explained by the very great extent of the atmospheres of these stars. (This has been shown directly for ε Aurigae.) Here the dimensions of the turbulence cells are without doubt small compared with the thickness of the reversing layer.

It follows from all that has been said that the problem of turbulence in stellar atmospheres is very complex. Besides the investigation of curves of growth, the line *contours* must be studied from every aspect. The latter are very important in connection with the question of the velocity distribution of turbulence cells. In Chapter 11 we took as a working hypothesis the law (11.32) for this distribution, v_0 being defined by formula (11.29). However, calculations show that various laws lead to very different contours. The curve of growth is thereby changed also. The value of v_t found by means of (11.32) may in some cases differ markedly from the actual mean turbulent velocity in the stellar atmosphere.

In conclusion, let us consider very hot stars, those of the O and early B classes. Here the situation is complicated by the fact that, besides the general true absorption, the scattering of radiation by free electrons may play a considerable part in the atmospheres of these

* The cells of small dimensions give the velocity $v_t = 5$ km/sec found from the curve of growth.

stars. Calculations which take account of the thermal motion of the electrons show that, for hot supergiants, the influence of this effect may be noticeable; the electron scattering increases the intensity of radiation in the central parts of the lines and decreases it in the wings.

4. **The Balmer series in the spectra of stars.** We now pass to the interpretation of the line contours of the Balmer series in the spectra of stars. These lines, as was stated at the end of Chapter 11, are usually broadened by interatomic electric fields. The broadening due to collisions does not play any part in this case. The problem of broadening by radiation damping can be settled by a comparison of the absorption coefficients given by formulae (11.39) and (11.56) for a fixed distance from the line centres $\Delta\lambda = \lambda - \lambda_0$. In the former, δ_{ik} is determined, according to (11.25), by the sum $\gamma_i + \gamma_k$, where $i = 2$ and $k = 3, 4, 5, \ldots$; here γ_2 and γ_k are given by formula (11.20). The value of γ_c can be taken as zero, since we are at present interested in radiation damping. Moreover, collision damping plays a negligibly small part compared with "statistical" broadening. Also, in formula (11.56) we must introduce for F_0 the expression (11.54). Here, in accordance with Chapter 11, n must be taken as the number of ions in 1 cm^3 (it is not yet certain what part electrons play; see Section 11.8). Since the number of ions in 1 cm^3 must be equal to the number of electrons in 1 cm^3, formula (11.56) can be rewritten

$$s_\lambda = (2{\cdot}61\, e)^{3/2} \frac{Cn_e}{(\lambda - \lambda_0)^{5/2}} = 321\, C\, \frac{p_e}{T} \frac{1}{(\lambda - \lambda_0)^{5/2}}, \tag{13.26}$$

where $\lambda - \lambda_0$ is expressed in ångströms. We recall that (13.26) holds only for the wings of the lines.

A comparison of (11.39) and (13.26) shows that, in practice, the broadening of the Balmer lines is determined in the majority of cases by formula (13.26). Thus, for example, for the line H_γ at a temperature of 10,000° and $\Delta\lambda = 4$ Å, radiation damping begins to play an important part when $p_e < 10$ bars; at $T = 5000°$ this pressure is halved. However, in many giants, and particularly in the supergiants, the broadening of the lines of the Balmer series must be determined mainly by radiation damping, and of course by the Doppler effect. These factors, as is shown by calculation, should play the principal part at $T_e > 30{,}000°$ not only in the atmospheres of supergiants, but also in those of stars of the main sequence. The broadening of lines by electric fields in the atmospheres of stars with $T_e > 30{,}000°$ should be important only for the higher terms of the series.

The idea that, in the majority of cases, the lines of the Balmer series, particularly its higher terms, are broadened by "microscopic" electric fields, and not by radiation damping, is confirmed by a number of facts, for instance:

(1) If the line widths of the Balmer series were determined only by radiation damping, then, by (13.25), the equivalent width W_λ would decrease very rapidly with increasing serial number k of the line. This follows because, for the first four lines of the Balmer series H_α, H_β, H_γ and H_δ, the oscillator strengths are respectively 0·637, 0·119, 0·0443, and 0·0212; for the high terms of the series the oscillator strength diminishes as k^{-3}. Also, when the number k increases, A_{ki} and λ decrease. Hence, if only radiation damping were present, even the H_γ line would have to be very narrow. The observations show, however, that in the spectra of early-type stars the equivalent width of the lines of the Balmer series *at first increases* as we pass from H_α to the higher terms, and then, having reached a distinct maximum, slowly decreases. In the spectra of the Sun and, in general, of stars similar to it in type, this maximum is considerably less pronounced.

These facts are easily explained in terms of the broadening of lines by electric fields. When the series number of the line increases, the number of split components of the upper term also increases (see Table 6), and in consequence the width of the line itself increases. Thus the observed relative constancy of the equivalent widths of lines of the Balmer series is a result of two oppositely acting effects: the decrease of the values of $f_{ik}\delta_{ik}$ and the increase of the splitting of the upper level when k increases. The decrease of the equivalent width when k increases further, after the maximum mentioned, arises because f_{ik} decreases, and consequently the optical depth in the central parts of the lines diminishes. Hence, from some k onwards, this optical depth will be less than unity, i. e. the lines are formed in *an optically thin layer*. In consequence, despite the continuing broadening (with increasing k) of the line wings [formula (11.55) and Table 6], the depth of the lines in their central parts diminishes, and this leads to a decrease in the equivalent width. Besides this, the overlapping of the line wings because of their close approach plays a part for high serial numbers. The law of variation of the coefficient of continuous absorption $\varkappa_\nu$ with frequency also plays a large part in the change of W_λ with serial number. This law is different for stars of different classes.

(2) The widths of the lines of the Balmer series in the spectra of dwarfs are considerably greater than in those of supergiants. Fig. 36

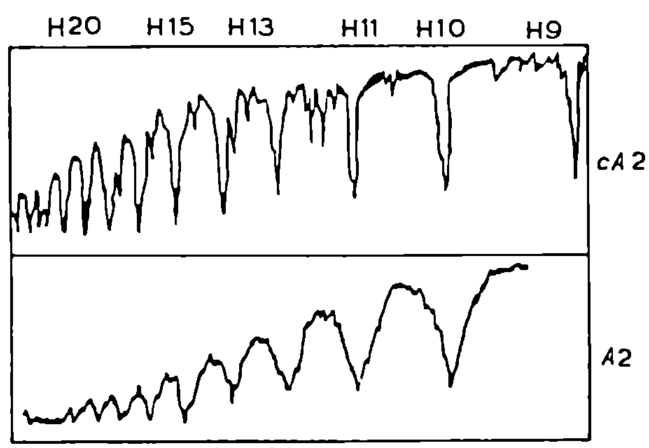

Fig. 36

serves to illustrate this effect. In it are reproduced microphotometer tracings of lines of the Balmer series for a supergiant of class cA 2 and an ordinary star of class A 2, obtained by G. A. Shaĭn from spectrograms taken at the 40 inch reflector of the Simeis observatory.

This effect is easily explained by means of formula (13.26). It follows from this formula that the broadening by electric fields is the greater, the greater p_e or n_e. Consequently the pressure effect will be less in the atmospheres of giants, where the total density, and therefore n_e, is smaller than in dwarfs.

Typical contours of the H_γ line, obtained from spectra taken with high dispersion, are shown in Fig. 37. The dotted line represents the contour of the H_γ line in the spectrum of the supergiant Rigel (class cB 8). The lower contour is that of the H_γ line in the spectrum of ζ Draconis, a main-sequence star of the same class.

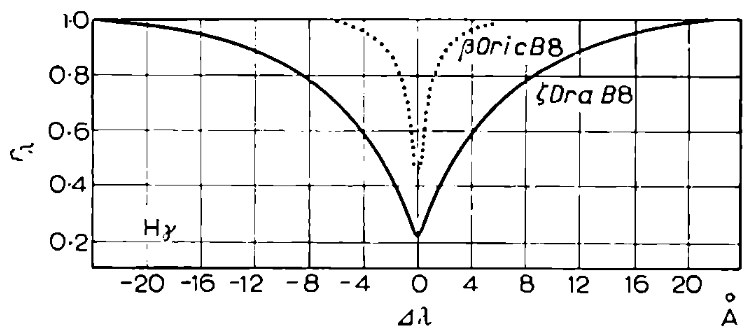

Fig. 37

It is interesting that the form of the contours of lines of the Balmer series in the spectra of B-type supergiants exhibits small changes with time. Moreover, these contours sometimes have a slight asymmetry, which also varies with time. Nothing like this is observed in the spectra of stars of relatively low luminosity (for example, the stars of the main sequence). This indicates a very unusual kind of structure in the atmospheres of supergiants.

White dwarfs also are characterised by a peculiar atmospheric structure. Because of the very large density of matter in these atmospheres, the lines of the Balmer series in the spectra of such stars are extremely broad.

Let us now consider the quantitative interpretation of the contours of the Balmer series in the spectra of stars. Here, as before, the chief problem is to find η_ν as a function of T, p, p_e and ϱ. Since the existing theory of the cores of the Balmer lines is very incomplete, we shall consider only the wings of the lines. Assuming the static case to hold, we can use formula (13.26).

The absorption coefficient s_λ, determined by formula (13.26), is referred to one neutral hydrogen atom in the second quantum state. Let there be $n_{0,2}$ such atoms in 1 cm^3. Then, from (13.2), (13.3) and (13.9), we find for the absorption coefficient referred to 1 gram of total mass

$$\sigma_\lambda = \frac{s_\lambda}{b\, m_H} \frac{n_{0,2}}{n_H}; \tag{13.27}$$

the ratio $n_{0,2} : n_H$ can be determined, T and p_e being known, by means of Saha's and Boltzmann's formulae (5.11) and (5.54). For $\varkappa_\nu$, the formula (13.17) must be used in the case of the solar atmosphere. In stars of class A 0, where the absorption is determined principally by neutral hydrogen, (13.16) must be used, and so on.

The quantity ε_ν needs special consideration. For the first few terms of the Balmer series, at not especially high temperatures, $\varepsilon_\nu \ll 1$. In this case, as before, ε_ν plays a considerable part only in the inner parts of the line, with which we are not at present concerned. However, for fairly high temperatures the density of radiation becomes so high that the majority of electrons which have made the transition $2 \to k$ will then be removed by radiation. This case corresponds to true selective absorption with $\varepsilon_\nu \approx 1$. The approximate equality $\varepsilon_\nu \approx 1$ holds for the higher terms of the Balmer series even at not especially high temperatures, since for these lines the binding energy $\chi_{0,k}$ is very small, and hence the probability of photo-ionisation from the level concerned is relatively high.

For the first few terms of the Balmer series in the spectra of fairly cool stars, we can take* $\varepsilon_\nu = 0$ without noticeable loss of accuracy.

The construction of contours of the first few terms of the Balmer series on the basis of formulae (13.26) and (13.27) has been performed by a number of authors, chiefly by means of the first two methods in Chapter 10. These calculations have led to the following conclusions. The general form of the theoretical contours, the widths of the lines, etc., are in satisfactory agreement with the results of observation. The best agreement between theory and observation is found for the outer parts of the line wings. This shows that the statistical broadening by electric fields is the principal factor in the broadening of the Balmer lines. However, a more detailed comparison of theory and observation reveals a number of important discrepancies. Thus, in the inner parts of the wings of the Balmer series in the spectra of hot supergiants, theory shows a steeper fall of r_ν to the centre of the line than is found from observation. In general, for the inner parts of the wings of the

* Some authors choose ε_ν by an experimental method, in which they seek to make the theoretical intensity [e. g. from formula (10.20)] in the centre of the line equal to the observed value.

Balmer lines, and particularly for the most central parts of these lines, the agreement between theory and observation is very often unsatisfactory. The cause of the discrepancy is not yet quite clear. However, we may notice some of the factors which affect the result. Firstly, formula (11.56) and all those derived from it are correct only for the purely static case where the perturbing particles (ions) are stationary with respect to the hydrogen atoms. Consequently, instead of formula (11.56) the more exact formula must be used which takes account of the relative motion of the perturbing and perturbed particles (see Chapter 11).

Furthermore, for the supergiants, the effect of turbulent motion may be very considerable; it influences most strongly the central parts of the lines.

Finally, it must be borne in mind that the Balmer lines are subordinate lines. Hence the non-coherence of the re-emission processes may be very important for these lines. This non-coherence should appear in two ways: firstly, the processes of ordinary scattering of the type $2 \to l \to 2$, where $l > 2$, are non-coherent (see Fig. 24); secondly, an electron which has made the transition $2 \to l$ may fall back not to the level 2, but to some other level, which may be the ground level. Similarly, an electron which has been raised to a higher level l not from the second level, but from some other, may then make the transition $l \to 2$, emitting a quantum of the Balmer series. Processes of the last two kinds (the interlocking of lines) lead, as it were, to an exchange of energy between spectral lines (see Chapter 14). These circumstances must be taken into account in the further development of the theory of Balmer line contours.

The problem of the central intensities in the lines of the Balmer series deserves special attention. We shall consider it in the next chapter. Here, however, it is necessary to consider one important circumstance which is directly related to the contours of the Balmer lines. Since the absorption coefficient is very large in the central parts of these lines, the observed radiation close to the frequency ν_0 will come from the most superficial layers of the stellar atmosphere, where the total density, and consequently p_e, is relatively small. Hence it may happen for sufficiently small p_e that s_ν is determined not by the broadening caused by the electric field, but mainly by radiation damping combined with the Doppler effect, i. e. by formula (11.39). But since the formation of the central parts of the lines is determined by the effect of a relatively small number of atoms, the contour in the line centre will be relatively narrow, in accordance with formula (11.39). Thus the contours of the Balmer lines will in this case have the following form: on a broad contour formed by the interatomic electric fields there should be in the centre a relatively narrow line produced by radiation damping

and the Doppler effect. It is clear that such a "composite" contour should be most distinctly shown in the spectra of ordinary stars, and not in those of supergiants, since in the latter case the broadening of the lines by the electric fields is very much less clearly seen than in the former.

G. A. SHAĬN [143] first called attention to the possibility of the appearance of these "composite" contours in the spectra of stars. He established that such contours are actually observed for the Balmer lines in the spectra of stars of fairly small absolute luminosity. The existence of such cores (of width of the order of 1 Å) in lines of the Balmer series is important because observers, in measuring the radial velocity of a star by means of hydrogen lines, determine the position of these same cores, which correspond to the highest layers of the atmosphere. By comparing the velocities thus obtained with those found from shallower lines (with smaller R_{ν_0}), which are formed at smaller heights, we can study the relative motion of gases in stellar atmospheres.

5. **The determination, from lines of the Balmer series, of the acceleration due to gravity in the atmospheres of stars.** When we have constructed theoretical contours of the Balmer lines for stars with various effective temperatures T_e and various accelerations due to gravity g at the surface, we can use these contours to determine the true value of g. To do this, we take some value for T_e which is close to the effective temperature of the given star, and compare the observed contour of some Balmer line with the sequence of theoretical contours for various values of g. The value of g which gives the best agreement between theory and practice is taken as the value sought. To obtain reliable results, the same must be done for other lines of the Balmer series also. Finally, a similar determination of the value of g can be performed in terms of the total absorptions (equivalent widths). For this purpose, the values of W_λ must first be calculated (again for a series of values of T_e and g) from the theoretical contours which have been constructed.

The important part which the Balmer series plays in the determination of g is due to two circumstances: firstly, the dependence of s_λ on n_e by formula (13.26) means that the dependence of the equivalent widths of the Balmer lines on the value of g is clearly marked*; giants give a considerably smaller W_λ and a contour completely different from that for dwarfs (see Figs. 36 and 37). Secondly, both observation and theory show that equivalent widths in the first few lines of the Balmer series

* For lines of metals, the radiation damping plays a more important part at low densities. For this reason the dependence of the equivalent widths on density is considerably less.

in the spectra of early-type stars depend much more markedly on g than on T_e. Hence the error in the T_e taken for the star does not cause a noticeable error in the value found for g.

In comparing the theoretical and observed contours, agreement should be sought for the outer parts of the wings, since in the inner parts, as we have said, the agreement between theory and observation is not satisfactory.

The method of determining g which we have discussed is of practical importance. If we assume that there is a one-to-one relation between the mass and the luminosity for the stars for which we determine g, then for a given T_e there is a one-to-one relation between g and the luminosity of the star. Thus the relation between W_λ and g is converted into one between W_λ and the star's luminosity. By this means the hydrogen lines can serve as a basis for determining spectroscopic parallaxes. Having found the star's luminosity from W_λ and knowing its apparent magnitude, we can determine the distance of the star.

The determination of g is equally important for supergiants. Calculations, both by the above method and by several other methods, have shown that the value of g found from absorption lines for these stars is considerably less than the ordinary "dynamical" value given by the expression

$$g = G\,\mathsf{M}/R^2\,, \tag{13.28}$$

where R is the radius of the star's photosphere. For this reason we shall call the value of g determined from spectroscopic data the **effective acceleration due to gravity** and denote it by g_{eff}.

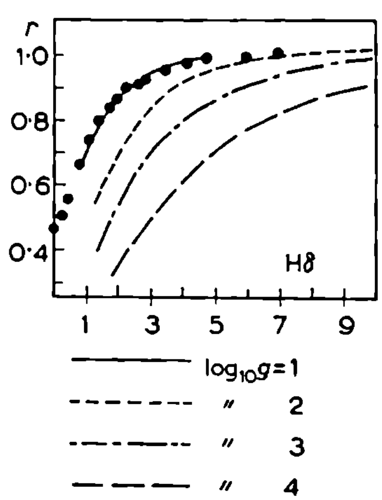

FIG. 38

The values of g_{eff} for supergiants of classes cB 5 to cA 3 have been determined by G. A. SHAĬN [139] by comparison of theoretical and observed contours. Fig. 38 shows, as an example, this comparison for the H_δ line in the spectrum of Rigel, a supergiant of class cB 8. It was found for seven supergiants of these classes which were studied that $\log_{10} g_{\text{eff}}$ lay between 1·0 and 1·4, and $\log_{10} g$ was about 3. The same estimate for $\log_{10} g_{\text{eff}}$ in the spectra of the supergiants studied by G. A. SHAĬN was obtained from many lines of other elements also. Such a small value of g_{eff} for the atmospheres of supergiants must, in the opinion of G. A. SHAĬN, be due to a very unusual atmospheric structure. These atmospheres must be extremely extended.

Similar results ($g_{\rm eff} \ll g$) are obtained for supergiants of other classes also. In particular, according to O. A. MEL'NIKOV (see note in Section 12.7), this inequality holds also for cepheids. In all these instances, the atmospheres of the stars are much more extended than would be the case for hydrostatic equilibrium with the value of g determined by formula (13.28).

However, the supposition that the small values of $g_{\rm eff}$ found from spectral lines are due only to the great extent of the atmospheres of the stars concerned involves some difficulties. In a recent paper [83], O. A. MEL'NIKOV has determined the *linear* thickness of the atmospheres of A-type stars of various luminosities. He finds that this thickness increases with the absolute luminosity of the star. When we go from stars with $M_v = +1^{\rm m}$ to those with $M_v = -5^{\rm m}$, the thickness increases by a factor of approximately 100. It is also found that the extent of the atmospheres of supergiants of these classes is insufficient to account for the small effective accelerations due to gravity $g_{\rm eff}$ that are found. Thus there are many points here which are still uncertain. It is therefore of particular interest to study the spectra of such supergiants as the cool K component of the system ζ Aurigae, where we can directly study (see Chapter 12) the distribution of density in the atmosphere of the K star. Here the rate of decrease of the density of matter is for metals 1000 times less than would be the case for hydrostatic equilibrium.

Despite the uncertainties mentioned above (the case of A-type supergiants), we can assert that some force directed away from the centre of the star, besides the gradient of the gas pressure, acts in the atmospheres of the supergiants. Some authors suppose that this force is the radiation pressure. It is more probable, however, that we are here concerned with forces similar to those which support the solar chromosphere (see Part III).

G. A. SHAÏN thinks that the extension of the atmospheres of supergiants is connected with the outflow of matter from them. This is confirmed by a study of the displacement of the lines of various elements in the spectra of supergiants. Furthermore, the extension of the atmospheres of supergiants must be related to the "turbulent" motions in these atmospheres. However, it is not yet entirely clear whether turbulence is the cause of the extension of the atmosphere.

6. Helium lines. The distortion of contours by neighbouring lines. Let us now turn to *helium* lines. The broadening of these lines is mainly determined by pressure effects. This is shown by a series of facts. The most important of these is the presence of forbidden helium absorption lines in the spectra of stars. For instance, in the short-wavelength

wing of the allowed line of neutral helium 4471·6 $2^3P^0 - 4^3D$, the "forbidden" line 4469·9 $2^3P^0 - 4^3F^0$ is observed. The appearance of forbidden lines is a result of the action of interatomic electric fields on the helium atoms. The intensity of the forbidden lines should increase with the field F_0, i. e. with the number n_e. Thus this intensity should be greater in the spectra of main-sequence stars than in the spectra of supergiants, where p_e is smaller. This is completely confirmed by the observations; see Fig. 39, where the contour of the lines 4471·6 Å and 4469·9 Å is shown for stars of different luminosities, but approximately the same spectral class. In the spectrum of the supergiant o^2 Canis Majoris, the forbidden line 4470 Å is absent, while in the spectrum of the main-sequence star γ Pegasi it is fairly strong. When g increases, the width of the line 4471·6 Å also increases, and this also is a result of the increasing pressure effect when p_e increases.

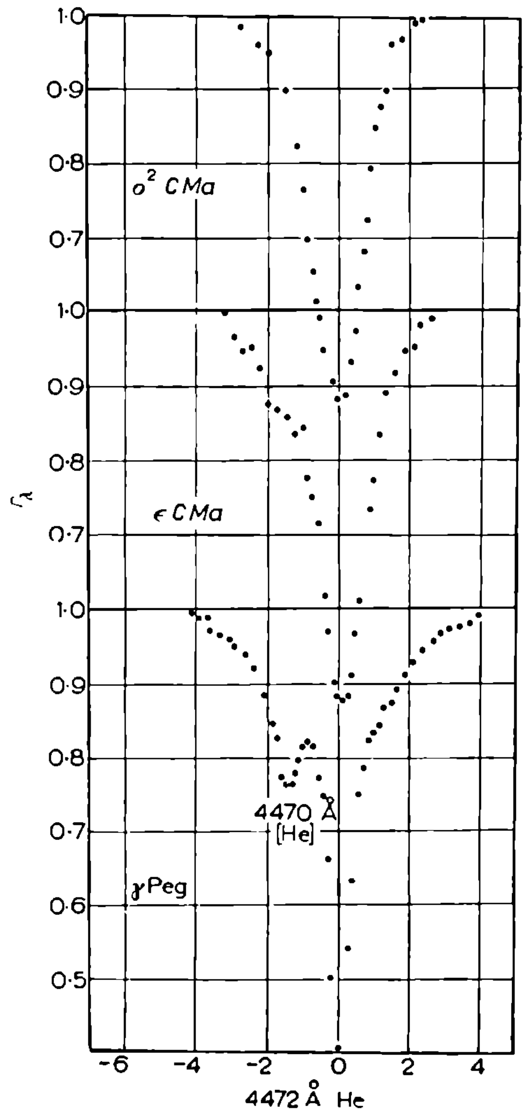

FIG. 39

Another fact which indicates the great importance of pressure effects in helium lines is that all lines of the diffuse subordinate series (P-D transitions) of *ortho-helium* and *parahelium* are broadened, while those of the sharp subordinate series (P-S transitions) are narrow. These results are related to the fact that the high D terms are split much more strongly by electric fields than the S and P terms.

The principal question arising in the interpretation of helium line contours is whether collision damping or interatomic electric fields determine the broadening of these lines. Existing investigations, both observational and theoretical, show that the chief factor is apparently collision damping (the perturbing particles must be mainly protons). In the broadening of helium lines by electric fields, the displacement of the split components should be asymmetric relative to the normal

position*. Thus in this case the absorption lines should also be asymmetric. However, observation shows that, for example, the lines 4472 and 4388 Å have an essentially symmetrical contour, as would be the case for collision broadening. Theoretical calculations also show that the width of the observed helium lines is determined by collision damping. Here the theoretical contours are slightly asymmetrical for fairly strong lines. This is due to the fact that statistical broadening by electric fields begins to appear in the wings, especially in their outer parts, and this must be greater, the greater is g. Apparently such "statistical" wings are to be seen for γ Pegasi in Fig. 39.

However, the quantitative investigation of helium line contours is as yet in an unsatisfactory state. In developing the theory, it is necessary to take account of Doppler broadening (low atomic weight of helium) and of radiation damping, as well as of pressure effects. Radiation damping is particularly important for the 2^1P^0 term, because of the frequent transitions of electrons to the ground level (this does not happen for the triplet system).

The lines of helium (P-D) for which the dependence of the splitting on the field intensity is not quadratic but linear, form a special case.

The equivalent widths of helium lines also exhibit a dependence on the absolute magnitude of the star, as may be seen, for instance, from Fig. 39. Hence helium lines also can be used to determine the value of g, though in this case the dependence of W_λ on g is much less than for hydrogen lines.

In concluding the discussion of hydrogen and helium lines, it is necessary to make the following remark. In calculating the effect of electric fields on the broadening of these lines, we are concerned only with interatomic "microscopic" electric fields. Electric fields of a "macroscopic" character which are large enough to produce a noticeable broadening and displacement of absorption lines have not yet been discovered in the atmospheres of stars, although the presence there of relatively small electromagnetic fields can scarcely be doubted. On the contrary, the presence of a *large* magnetic field (up to 5000 gauss) on the surfaces of some stars has been established beyond all doubt by H. W. Babcock and others. In many of these stars the magnetic field is variable, changing sign periodically, and these changes are accompanied by variations in the intensities of some absorption lines. The question of the nature of the general magnetic fields on the surfaces of stars is so far unsettled. The most recent survey of this topic is by H. W. Babcock and T. G. Cowling [14].

---

* This is the case in general with the quadratic, but not with the linear, Stark effect.

The general ideas of the theory of absorption-line contours which we have considered have related to stars in a steady state, and the broadening of the lines was determined mainly by the properties of atoms and the interaction between them (pressure effects). Besides these, there are other factors which lead to distortion of the "normal" contours. One of these ("turbulence") we have considered. As well as being the effect of turbulence, distortion of absorption-line contours is produced by neighbouring absorption lines, and also by the rotation of stars and the outflow of matter from them. However, the latter effect is large only for non-steady stars and in general for stars with spectra of unusual character, which we do not consider here.

We shall make a number of remarks relating to the distortion of contours by neighbouring absorption lines. In some cases (especially for late-type stars) this effect is very considerable. In particular, in the parts of the spectrum (chiefly in the ultra-violet) where there is a strong concentration of lines, the mutual overlapping of the wings of neighbouring lines not only distorts the lines themselves, but may produce a general weakening of the continuous spectrum in these regions.

In studying the contour of a line (1) which is distorted by another line (2), it is necessary to begin from the general equation of transfer (9.29). Here it must be remembered that the absorption coefficient σ_ν in the given frequency, by virtue of the additive nature of the energy absorbed, is the sum of the coefficients $\sigma_{\nu,1}$ and $\sigma_{\nu,2}$ relating to the two lines. This, of course, introduces a complication into the theory, since the quantities $\eta_{\nu,1}$ and $\eta_{\nu,2}$ may vary with frequency and depth in different ways. It must also be borne in mind that, if the distorting line (2) is formed as a result of true selective absorption processes (i. e. $\varepsilon_\nu = 1$ for it), then $\sigma_{\nu,2}$ plays the same part as the coefficient of continuous absorption.

If, for example, in the wing of the stronger line (2) a faint line (1) is observed, this means that the faint line arises in relatively superficial layers of the stellar atmosphere, since the wing of the line (2) prevents any radiation from emerging from the lower layers of the atmosphere. By comparing such lines with lines which are not weakened by other lines, we can draw a number of important conclusions concerning the physical state (and chemical composition) of various layers of the stellar atmosphere.

The absorption of radiation beyond the limit of the Balmer series plays a similar part in the spectra of early-type stars. Here the coefficient of continuous absorption is so large that we are observing radiation coming from practically the outermost layers of the star. Hence the absorption lines in this spectral region will be weakened. From a comparison of absorption lines lying up to and beyond the limit of the

Balmer series, we can again draw a number of important conclusions concerning the physical state (and chemical composition) of various layers in stellar atmospheres. However, this subject has so far not been much developed.

7. **The rotation of stars.** To conclude the present chapter, let us consider the distortion of absorption-line contours owing to the rotation of stars. This rotation causes a broadening of all the absorption lines in a star's spectrum, since the different parts of the visible surface of a rotating star move with different velocities relative to the observer. Hence the absorption line for each element of the star's visible disc will be displaced by a definite amount owing to the Doppler effect. On observing the *whole* disc of the star, we find a broadening of the line. The greatest rates of rotation are encountered among early-type stars. For binary systems, the rates of rotation of the component stars are statistically the greater, the shorter the period of revolution of the system and the greater the radial-velocity amplitude K of the system.

The main problem of the theory of absorption lines distorted by rotation is the comparison of theoretical contours with observation in order to determine the rate of rotation of the star. This problem was first solved by the work of G. A. Shaĭn and O. Struve [147].

The following method has been the one most used to determine the rate of rotation of a star from a comparison of the theoretical and observed contours. It is a development of the theory evolved by G. A. Shaĭn and O. Struve. To construct the contour of any line in the spectrum of a rotating star, we take, as an initial contour, the observed contour of the same line in a non-rotating star of the same class. Then, giving various values to the equatorial velocity of rotation of the star, we construct for each of these values the theoretical contour distorted by rotation. By comparing the observed contour of the same line in the spectrum of a rotating star (of the same class) with the constructed sequence of theoretical contours, we find the speed of rotation v of the star, or, more precisely, the value of $v \sin i$, where i is the angle between the line of sight to the star and its axis of rotation (see below).

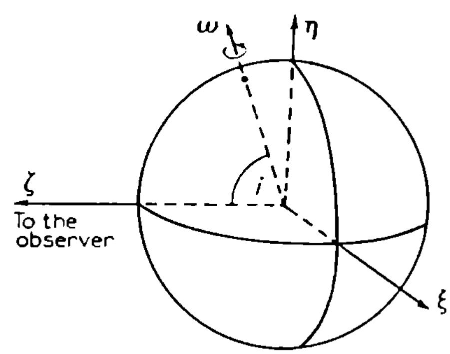

Fig. 40

In connection with the above we now consider how we can pass from the contour in a non-rotating star to the contour distorted by rotation. We place the origin of a system of rectangular co-ordinates ξ, η, ζ at the centre of the star, with the ζ axis towards the observer (Fig. 40). The angular velocity vector ω is, as usual, directed along the

axis of rotation of the star, while the η axis is so placed that the vector $\boldsymbol{\omega}$ lies in the $\zeta\,\eta$ plane. In this case, the vector $\boldsymbol{\omega}$ has the components

$$\boldsymbol{\omega} = \{0\,,\,\omega \sin i,\,\omega \cos i\}\,, \tag{13.29}$$

where, as was said above, i is the angle between the line of sight to the star and its axis of rotation.

The position of any point on the surface of the star is defined by the vector

$$\mathbf{r} = \{\xi,\,\eta,\,\zeta\}\,. \tag{13.30}$$

Also, the linear velocity vector of this point is

$$\mathbf{v} = \boldsymbol{\omega}\wedge\mathbf{r}\,. \tag{13.31}$$

Consequently, the component of this velocity in the line of sight, i. e. along the ζ axis, is

$$v_\zeta = \omega_\xi\,\eta - \omega_\eta\,\xi = -\,\xi\,\omega \sin i\,, \tag{13.32}$$

since $\omega_\eta = \omega \sin i$, and $\omega_\xi = 0$. We see from (13.32) that the radial velocity of the separate points on the visible disc of the star depends only on the co-ordinate ξ. Consequently, each element of an infinitely narrow strip, parallel to the η axis, on the visible disc of the star has the same radial velocity.

According to (11.30) and (13.32), each point of a contour corresponding to a strip between ξ and $\xi + d\xi$ is displaced from its position in a contour not distorted by rotation by an amount whose absolute value is

$$\Delta\lambda = \lambda - \lambda_0 = \frac{\lambda}{c}\,\xi\,\omega \sin i\,. \tag{13.33}$$

If, further, R is the radius of the star, then (13.33) can be rewritten

$$\Delta\lambda = \lambda - \lambda_0 = \frac{\lambda}{c}\,\frac{\xi}{R}\,v \sin i\,, \tag{13.34}$$

where $v = \omega R$ is the linear velocity of rotation of the star's atmosphere at the equator. Here $v \sin i$ is the radial velocity at the equator for the points $\xi = \pm R$.

Let us now obtain an expression for the intensity of radiation at any point of a contour distorted by rotation. We shall take the radius of the star as unity; the extreme values of ξ and η will then be ± 1. Let $I(\xi,\,\eta,\,\lambda - \lambda_0)$ be the intensity of radiation on the disc of a non-rotating star at a point with co-ordinates ξ and η and at a distance $\lambda - \lambda_0$ from the centre of the line. If the star is rotating, the whole line contour at the point $(\xi,\,\eta)$ is displaced by an amount $\Delta\lambda$, determined by formula (13.34).

We must now introduce $\lambda_0 \pm \Delta\lambda$ into the expression for the intensity I in place of λ_0. The energy sent out to the observer by the element $d\xi\, d\eta$ of the stellar disc in 1 second inside unit solid angle, in the wavelength interval from λ to $\lambda + d\lambda$, is

$$I\left(\xi,\, \eta,\, \lambda - \lambda_0 \left[\, 1 \pm \xi\, \frac{v}{c} \sin i\, \right]\right) d\xi\, d\eta\, d\lambda\,. \tag{13.35}$$

If the intensity of the continuous radiation were constant over the entire disc of the star, and the line contour had the same form at every point of the disc, then, by (13.35), the value of I for a given ξ would be independent of η. This condition is not fulfilled for actual stars.

In order to obtain the energy emitted by the whole disc in the given wavelength, we must integrate (13.35). As a result, we obtain for the energy referred to unit wavelength interval

$$E_\lambda = 2 \int\limits_{-1}^{1} \left\{ \int\limits_{0}^{\sqrt{(1-\xi^2)}} I\left(\xi,\, \eta,\, \lambda - \lambda_0 \left[\, 1 \pm \frac{\xi}{c}\, v \sin i\right]\right) d\eta \right\} d\xi\,. \tag{13.36}$$

The calculation of E_λ for given v sin i and λ is performed as follows. We divide the visible disc of the star into a series of narrow strips parallel to the η axis, and for each of these strips (of width $\Delta\xi$) we calculate the value of

$$\left\{ \int\limits_{0}^{\sqrt{(1-\xi^2)}} I\, d\eta\, \right\} \Delta\xi\,.$$

The summation of all such elements is then effected. To calculate the quantity in braces in (13.36) we use the intensities $I(\xi,\eta)$ for the case of a non-rotating* star, but for a given λ we take the point in the contour which corresponds to a distance $\lambda - \lambda_0\, [1 \pm (\xi/c)\, v \sin i]$ from the centre of the line. We carry out a similar procedure for other λ also, and as a result the line contour distorted by rotation is found.

In calculating E_λ from formula (13.36) it must be borne in mind that in general there is a darkening to the limb of the star's disc in every wavelength. We shall assume that the line contour is the same over the whole disc of the star, i. e. that the dependence of r_λ on λ is the same at all points of the disc. In this case the initial intensity distribution within the line for any point ξ, η of the stellar disc, i. e. the function $I(\xi,\, \eta,\, \lambda - \lambda_0)$, is

$$I(\xi, \eta, \lambda - \lambda_0) = \Phi(\lambda, \theta)\, I(0, 0, \lambda - \lambda_0)\,, \tag{13.37}$$

* The case where the gravitational effect (see below) is important forms an exception. Here the figure 2 before the integral in (13.36) must be removed, and the integration over η is from $-\sqrt{(1-\xi^2)}$ to $\sqrt{(1-\xi^2)}$.

where $\Phi(\lambda, \theta)$ is the law of darkening of the star's disc for the given wavelength, $\sin\,\theta$ being $\surd\,(\xi^2+\eta^2)$. The function $\Phi(\lambda, \theta) \approx \Phi(\lambda_0, \theta)$ within the line is practically constant. For some eclipsing variables it can be determined from observation. The appropriate calculations show that the effect of limb-darkening is generally small.

If the contour itself changes with distance from the centre of the disc, then the calculations are more laborious. However, it has not hitherto been necessary to take this factor into account.

At very high velocities of rotation the **gravitational effect** has to be taken into account; it operates as follows. The general theory of rotating stars shows that the flux πH_ν at the surface of a star which rotates as a rigid body is proportional to the effective acceleration due to gravity at the point. Consequently, the effective temperature at the poles must be greater than at the equator. Thus, the star's rotation in this case alters the energy distribution over the star's disc. Since in general the angle i is neither 0° nor 90°, there will be an asymmetry in the energy distribution over the star's disc, relative to the $\xi\zeta$ plane. However, the gravitational effect causes an appreciable distortion of line contours only for velocities of rotation so large that the centrifugal force at the equator is of the same order of magnitude as the force of gravity. Finally, the differential effect of rotation may play a considerable part in the case of rotating stars, since the outer layers of the star may have different velocities at different latitudes. This is so in the case of the Sun, as is well known.

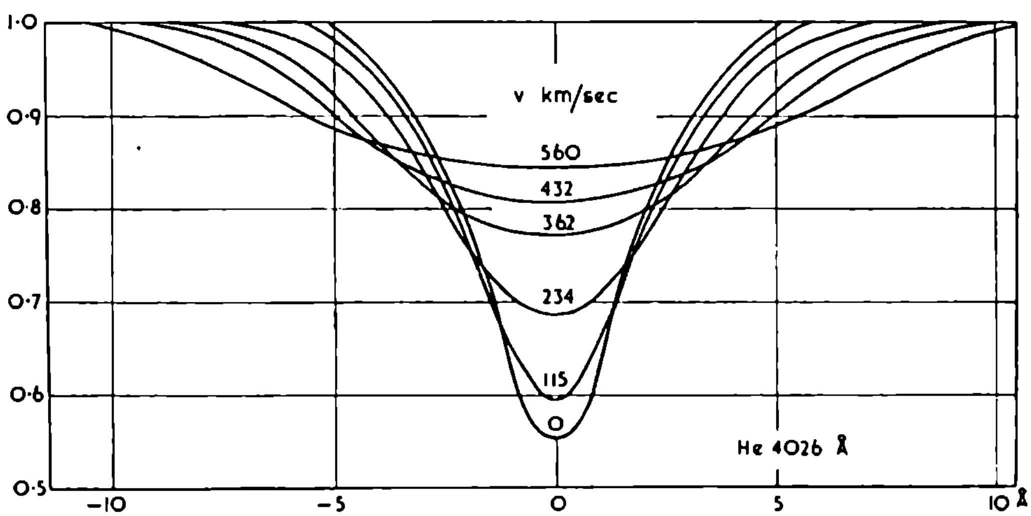

Fig. 41

As an application of formula (13.36), we give in Fig. 41 the contour of the line He I 4026 Å as distorted by rotation. The contour with $v = 0$ is the contour of the line 4026 Å in the spectrum of the non-

rotating star ι Herculis. The theoretical contours are calculated for the case where limb-darkening is absent. To take account of this effect would cause little further alteration; it would make the contours a little deeper and less broad. We see that, as $v \sin i$ increases, the contour is raised and broadened, while the equivalent width of the line is not affected by the star's rotation.

Graphs of the type shown in Fig. 41 form the starting-point for the determination of $v \sin i$ from observed contours. To compare the theoretical and measu redcontours, they must be reduced to the same equivalent width. It must be emphasised that by the method described we determine not v_{eq} but $(v \sin i)_{eq}$. Finally, it must be stated that, in the method of determining $(v \sin i)_{eq}$ which we have explained, it is not necessary to take account of instrumental distortion if the original contour and the observed contours, distorted by rotation, are obtained with the same instrument. It is evident that the effect of rotation operates more strongly on the deeper lines than on those which are shallower and more obliterated. This is one of the criteria which enable us to decide whether the star is rotating or not. Besides this there are other criteria. For example, the rotation of a star causes a broadening of all the absorption lines. Also, the broadening of the lines in the spectrum of a rotating star increases, roughly speaking, proportionally to the wavelength [see (13.34)]. Finally, rotation, unlike many other effects, has no influence on the ratios of the equivalent widths.

At the present time a fairly large amount of material has been accumulated on the values of $v \sin i$. It has been ascertained that in some cases these values can be very large. Thus, among stars of class B, we find some for which $v \sin i$ reaches 400 km/sec. Stars of class B for which $v \sin i$ exceeds 200 to 300 km/sec are not infrequently met with. The greatest velocities of rotation belong to stars of class Be. For some of these, $v \sin i$ exceeds 500 km/sec. Stars of the later classes rotate with considerably smaller velocities. Thus, the linear velocity of rotation of the Sun at the equator is only 2 km/sec. Cases of rapid rotation in stars of late spectral classes are exceedingly rare. They occur mainly among stars of W Ursae Majoris type, which have shallow lines in their spectra, and these are certainly distorted strongly by rotation. The G-type star HD 117555 also should perhaps be regarded as a rapidly rotating star. The H_α line in its spectrum is bright. The equatorial velocity of rotation for this star is 75 km/sec.

The accumulation of reliable data on the values of $v \sin i$, for as many rotating stars as possible, is a very important task of theoretical and observational stellar spectroscopy.

In conclusion, we may remark that in the spectra of eclipsing stars an asymmetric distortion of the lines is observed; this is fully explained

by the rotation of the eclipsed star of the system. When we see behind the eclipsing star one of the limbs of the disc of the rotating eclipsed star, the approach (or recession) of this limb to (from) the observer displaces the absorption line in only one direction, and since we see a finite portion of the disc, corresponding to various $\xi v \sin i$, the resulting line will be asymmetrical.

Chapter 14. The variation of absorption-line contours from the centre to the limb of the solar disc. Non-coherent scattering. Central residual intensities

1. The variation of absorption-line contours from the centre to the limb. Interlocking. The analysis of absorption-line contours *at different points on the solar disc* is the principal method of testing the theory of absorption lines in the solar spectrum. For in this case we are examining *in different directions* the layers which form the lines, i. e. we are introducing a new variable θ into the problem. By this means we can test whether we have found the correct distribution with depth of the various characteristics of the atmosphere. These include the form of the interaction between radiation and atoms. For example, by studying line contours at different points of the solar disc, we can find which process (scattering or true absorption) plays the chief part in line formation at various levels. By this means, for instance, we were able to establish at the beginning of Chapter 9 that in the solar atmosphere strong lines, at least, are formed by scattering of light.

Let us assume that, starting from definite hypotheses and using parameters chosen in the proper manner, we have obtained complete agreement between the theoretical and observed contours at the *centre* of the solar disc. Then a comparison of the results of theory and observation for *various points of the disc* is a test of the original theoretical assumptions and parameters. This method is also applicable to eclipsing variables. Here the lines in the spectrum of the eclipsed star both during and outside eclipse must be studied.

In order to study the contour of any absorption line at various points of the disc, we must solve the equation of transfer (10.2) or (9.32). Equation (10.2) can be rewritten

$$\begin{aligned} \mathrm{d}I_\nu(\theta,\tau_\nu)/\mathrm{d}\tau_\nu - I_\nu(\theta,\tau_\nu)\,(1+\eta_\nu)\sec\theta\,+ \\ +\,[J_\nu\,\eta_\nu(1-\varepsilon_\nu) + B_\nu(1+\eta_\nu\,\varepsilon_\nu)]\sec\theta = 0\,, \end{aligned} \tag{14.1}$$

i. e. we have here an equation of the form (3.20). According to (3.21), the solution of equation (14.1) can be written down as

$$I_\nu(\theta,\tau_\nu) = A_\nu \exp\left[\sec\theta \int_0^{\tau_\nu} (1+\eta_\nu)\,\mathrm{d}t_\nu\right] + $$
$$+ \exp\left[\sec\theta \int_0^{\tau_\nu} (1+\eta_\nu)\mathrm{d}t_\nu\right] \int_{\tau_\nu}^{\infty} [(1-\varepsilon_\nu)\,\eta_\nu J_\nu + \tag{14.2}$$
$$+ (1+\varepsilon_\nu\eta_\nu)\,B_\nu] \exp\left[-\sec\theta \int_0^{t_\nu} (1+\eta_\nu)\,\mathrm{d}t_\nu\right] \sec\theta\,\mathrm{d}t_\nu\,,$$

where A_ν is a constant of integration. Proceeding as in the derivation of (3.36) and (3.37), we find that $A_\nu = 0$, and consequently the emergent intensity $I_\nu(\theta,0)$ at the boundary of the atmosphere is

$$I_\nu(\theta,0) = \int_0^{\infty} [(1-\varepsilon_\nu)\,\eta_\nu J_\nu + $$
$$+ (1+\varepsilon_\nu\,\eta_\nu)\,B_\nu] \exp\left[-\sec\theta \int_0^{t_\nu} (1+\eta_\nu)\,\mathrm{d}t_\nu\right] \sec\theta\,\mathrm{d}t_\nu\,. \tag{14.3}$$

Denoting the corresponding intensity in the continuous spectrum in the frequency ν by $I_\nu^0(\theta,0)$ and putting in this case $\eta_\nu = 0$, we find for the residual intensity $r_\nu(\theta)$, at the point of the disc considered, the expression

$$r_\nu(\theta) = I_\nu(\theta,0)/I_\nu^0(\theta,0)$$
$$= \frac{\int_0^{\infty} [(1-\varepsilon_\nu)\,\eta_\nu J_\nu + (1+\varepsilon_\nu\,\eta_\nu)\,B_\nu] \exp\left[-\sec\theta \int_0^{t_\nu} (1+\eta_\nu)\,\mathrm{d}t_\nu\right] \sec\theta\,\mathrm{d}t_\nu}{\int_0^{\infty} B_\nu\, e^{-t_\nu \sec\theta} \sec\theta\,\mathrm{d}t_\nu}\,. \tag{14.4}$$

Thus, in order to find the dependence of $r_\nu(\theta)$ on the frequency for various given θ, we need to know the dependence of the quantities B_ν, J_ν, η_ν and ε_ν on the optical depth. We have already discussed in detail, in Chapters 10 and 13, how this is to be found for B_ν, J_ν and η_ν. The dependence of ε_ν on the depth will be considered at the end of the present chapter. In the majority of cases the integrals in (14.4) are calculated by the method of numerical integration. However, in some cases the value of $r_\nu(\theta)$ can be obtained in closed form. For example, for the method explained in Section 10.2, the functions J_ν and B_ν can be taken from formulae (10.12) and (8.10) respectively, where C_ν

is determined by the expression (10.17), and the parameters η_ν and ε_ν are constant throughout the photosphere. Effecting the elementary integration, we obtain for $r_\nu(\theta)$

$$r_\nu(\theta) = [a_\nu + b_\nu \cos\theta]^{-1} \left\{ a_\nu + \frac{b_\nu \cos\theta}{1+\eta_\nu} + \right. \\ \left. + \frac{\eta_\nu(1-\varepsilon_\nu)\left[\frac{3}{3} b_\nu - a_\nu(1+\eta_\nu)\right]}{[1+\eta_\nu+q_\nu\cos\theta]\left[\frac{3}{3} q_\nu + 1 + \eta_\nu\right]} \right\}, \tag{14.5}$$

where q_ν is given by the expression (10.9).

The expression (14.5) for $r_\nu(\theta)$ can be considered separately for *scattering* alone ($\varepsilon_\nu = 0$) and *absorption* alone ($\varepsilon_\nu = 1$) and the resulting contours compared. Such a comparison shows that, within the *wings* of the lines, the residual intensity (for a definite wavelength in the line) and its law of variation from the centre to the limb of the solar disc are almost independent of the value of ε_ν, i. e. of whether the absorption line (or more precisely its wings) is formed by scattering of radiation or by true absorption.

It is also of interest to compare the dependence of $r_\nu(\theta)$ on the angle θ for two models, that just considered and the simple model of Chapter 9, in which the star's atmosphere is divided into the "reversing layer" and the "photosphere". For *scattering*, the value of $r_\nu(\theta)$ is given by formula (9.24); for *absorption*, the corresponding formula can be derived by separating off some region between the limits $\tau_\nu = 0$ and $\tau_\nu = \tau_{\nu,1}$ in the photospheric models of Chapter 3. The absorption lines are to be formed in this region.

A comparison of these models shows that the variation of the intensity of the wings in passing from the centre to the limb of the disc of the star (or of the Sun) is very different in the two cases, this result being true both for a scattering mechanism and for an absorption mechanism. In other words, the variation of the contour in the line wings in passing from the centre to the limb of the disc depends markedly on the structure of the stellar atmosphere.

Finally, in the central parts of fairly strong lines, the variation of $r_{\nu_0}(\theta)$ with the angle θ should depend very much on what is the main process for these frequencies, absorption or scattering. Thus, for instance, in the first case $I_{\nu_0}(\theta)$ is close to $B_\nu(T_0)$ for all θ and depends only slightly on θ, while in the case of scattering the intensity $I_{\nu_0}(\theta)$ varies quite differently with θ.

Let us consider briefly some of the principal observational results regarding the law of variation of line contours over the solar disc. For *strong* lines the most usual relation is that, at the extreme limb of the disc (in practice, for $R \approx 0{\cdot}995\ R_\odot$), the values of $r_\nu(\theta)$ throughout the contour are greater than the corresponding $r_\nu(\theta)$ at the centre of

the disc (cf. Figs. 23 and 35). In general, the variation of the contour over the whole disc (from $R = 0$ to $R = 0{\cdot}995\ R_\odot$) is often very complex and is not the same for different lines.

The equivalent widths of *faint* lines, whose contours are instrumental in nature, usually increase towards the limb, but at a certain distance from the centre of the disc a maximum is reached, and then the equivalent width decreases again. For *very faint* lines, the increase of W_λ is apparently maintained up to the limb itself. In this respect very faint lines are sharply distinguished from strong lines, where W_λ is less at the limb than at the centre. The central residual intensities of faint lines, unlike those of strong lines, are practically the same over the whole disc of the Sun.

Let us now examine the results of quantitative comparisons between the general theory which has been explained above and the observations. In Fig. 35 we have already given an instance of such a comparison, from which a satisfactory agreement is seen to exist between the theoretical and observed contours, although there are also definite discrepancies. The conclusion that there are definite discrepancies between theory and observation follows also from the majority of other work on the variation of contours over the solar disc*. This divergence is found both for strong lines (particularly in the ultra-violet region of the spectrum) and for faint lines.

The cause of these discrepancies might be supposed to lie in the insufficiently exact calculation of the integrals on the right of (14.4); rejecting the various simplifications such as, for example, $\eta_\nu =$ constant, $B_\nu = a_\nu + b_\nu t_\nu$, etc., we might attempt to refine the methods of calculation of these integrals. In particular, the direct method discussed at the end of Chapter 7 might be used to find the relation between B_ν and t_ν. Furthermore, the variation of $I_\nu(\theta, \tau_\nu)$ with θ can be taken into account fairly exactly in deriving the dependence of J_ν on τ_ν, and so on. However, all these improved calculations, though few have been performed as yet, show that the use of the expression (14.4) cannot give a sufficiently satisfactory agreement between theory and observation. It is true that the existing observations still need further refinement. The results obtained by different observers are usually somewhat discordant. However, the discrepancies between theory and observation exceed the possible errors in the latter.

Faint lines form a special case. Observations during eclipses have established that these lines are observed as absorption lines at points on the extreme limb of the solar disc, where they ought to change into emission lines (see Chapter 21). This fact could be explained by

* See [168, § 113] for a detailed discussion of such investigations and of the discrepancies found.

(e. g.) an unevenness of the Sun's surface, owing to which we observe radiation emerging from parts of the solar surface which are oriented somewhat differently with respect to the observer. Consequently, even where, for a "smooth" surface, we should expect $\theta \approx \frac{1}{2}\pi$, there are regions with an effective θ differing from this value. This, of course, reinforces the absorption lines at the limb itself.

Let us return to *strong* lines. The impossibility of reconciling the results of applying the very general formula (14.4) with observation compels us to conclude that the initial assumptions behind this formula are not completely accurate. We recall that this formula follows directly from the equation of transfer (9.29). At the same time, the principal assumption in this equation of transfer is that the re-emission within absorption lines takes place without change of frequency. In other words, each absorbed quantum is re-emitted in the same frequency. Let us now consider how far this assumption is justified.

We first consider re-emission processes for transitions whose lower level is not the ground level. The importance of this case lies in the fact that in such transitions subordinate lines are formed, and there are many more of these lines in stellar spectra than there are resonance lines. The formation of subordinate lines differs from that of resonance lines, which start from the ground level, in the following ways (see Fig. 24):

(1) After the absorption of a quantum by an atom as a result of the transition $k \to l$, the electron may fall to a quite different level, say the ground level, where $i = 1$. In this case, the frequency of the re-emitted quantum will be totally different from that of the absorbed quantum, and the radiation corresponding to the transition $k \to l$ (i.e. to the coefficient j_ν) will be attenuated. However, transitions in the opposite direction, of the type $1 \to l \to k$, will contribute to the strengthening of the re-emission in the frequencies corresponding to the transition $k \to l$. Thus the energy balance in a subordinate line ($k \to l$ transitions) can be computed only by taking account of the relation between this line and other lines having the upper or lower level in common with the given line. In what follows we shall call this the **interlocking effect**.

(2) The lower level of a subordinate line is not sharp. Hence, after a quantum $h\nu'$ has been absorbed by the atom, the electron which has made the transition $k \to l$ can fall back to the same level k but to a different (energy) element of this level, as a result of which a quantum $h\nu''$ will be emitted which differs slightly from the absorbed quantum $h\nu'$. In this case we shall call the re-emission process **non-coherent**, bearing in mind that the concept of non-coherence here refers to the frequency, and not to the phase. It is clear that the phase remains the same during the re-emission process considered (in the absence of such factors as, for instance, collisions).

The interlocking effect has been studied by a number of authors. It may be somewhat different in character for each subordinate line, since the scheme of lines connected with different lines must in general be different. Without taking account of non-coherence in frequency, the interlocking effect for any line can be investigated on the basis of

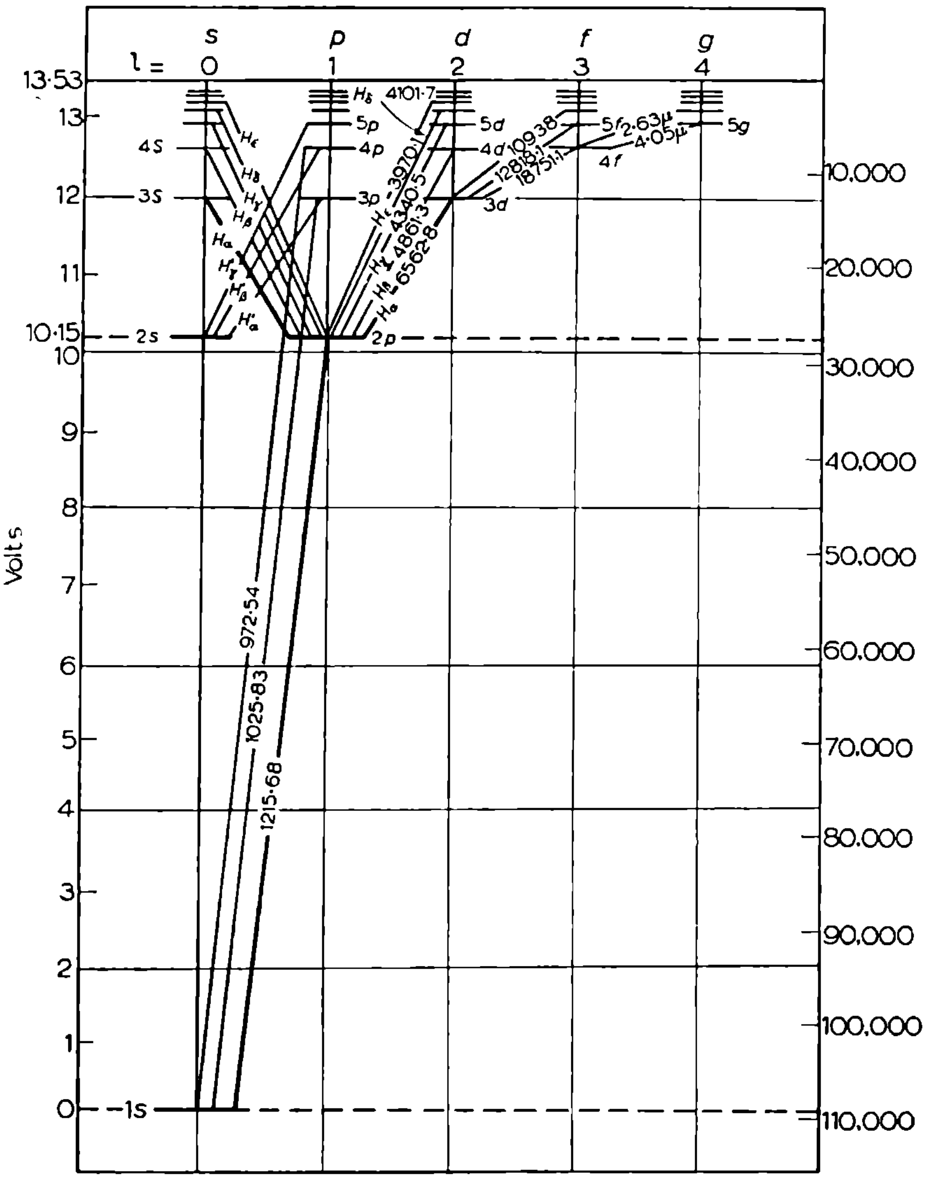

Fig. 42

the balance between the number of transitions in the given line, allowing for the most important lines connected with it. This balance makes it possible to calculate the ratio of the numbers of atoms in the upper and lower states, and thereby the variation of j_ν due to the interlocking effect. As an example of the calculation of interlocking (neglecting non-coherence), let us consider one of the most important lines in stellar

spectra, the H_α line, which begins from the second level. It may be assumed that the preponderant part of the absorption processes leading to the formation of the H_α line correspond to the transitions $2\,p \rightarrow 3\,s$ and $2\,p \rightarrow 3\,d$. The $2\,s \rightarrow 3\,p$ transitions are few in number compared with these (Fig. 42). On the other hand, having arrived at the $3\,s$ or $3\,d$ level, the electron will usually return to the $2\,p$ level again, since the probability of upward transitions from the $3\,s$ and $3\,d$ levels is small, for fairly cool stars, compared with the probability of some transition $3 \rightarrow 2$. The transitions $3\,s \rightarrow 1\,s$ and $3\,d \rightarrow 1\,s$ are forbidden by the selection rules for the azimuthal quantum number. Thus, the interlocking effect for the case considered (the H_α line) is connected only with the processes $2\,s \rightarrow 3\,p$ (which can lead on to the transition $3\,p \rightarrow 1\,s$, i. e. the emission of a Lyman β quantum). Hence, on account of the relative rarity of $2\,s \rightarrow 3\,p$ transitions, the effect considered is not so large as might appear at first sight.

2. **Non-coherence of scattering processes.** We have now considered the interlocking effect for a line of the Balmer series and have shown it to be small. However, in other cases it may be more considerable. We must simultaneously take into account the effect of non-coherence, as well as of interlocking.

For *non-coherent* scattering of light, the fundamental expression (9.5) for the emission coefficient j_ν

$$j_\nu = \sigma_\nu \int I_\nu \frac{\mathrm{d}\omega}{4\,\pi} \tag{14.6}$$

must evidently be replaced by the following [155]:

$$j_\nu = \iint I_{\nu'}\; p\,(\nu, \nu', \gamma)\, \sigma_{\nu'}\, \mathrm{d}\nu'\, \frac{\mathrm{d}\omega}{4\,\pi}\,, \tag{14.7}$$

where $p(\nu, \nu', \gamma)\,\mathrm{d}\nu$ is the probability that radiation of frequency ν' is scattered into the frequency interval from ν to $\nu + \mathrm{d}\nu$, in a direction making an angle γ with the original ray.

The importance of non-coherent scattering processes in stellar atmospheres has been discovered only very recently. In particular, it has been found that non-coherence (in both frequency and phase) exists for resonance lines also. Let us consider the reasons why non-coherence, both for subordinate and for resonance lines, may occur.

(1) Non-coherence arising from thermal, turbulent and other motions of the absorbing and emitting atoms. In general, the components of the velocity of the moving atom in the direction of motion of the absorbed and emitted quanta are not the same. Hence the frequencies of the absorbed and emitted quanta must in general be different, even if the scattering by stationary atoms of the kind in question is completely coherent.

(2) Non-coherence due to the broadening of the lower level. This applies mainly to subordinate lines*. Here the magnitude of the re-emitted quantum $h\nu''$ is in general (see Fig. 24) different from that of the absorbed quantum $h\nu'$. If a strong radiation field is present (very hot stars), the ground level also may be broadened [see formula (11.20)]. In this case the coherence is destroyed for resonance lines also.

It is of interest to remark that, in the case we are now considering, the following restrictive condition on processes of emission and absorption of quanta exists: in the given line, the sum of the values of $h\nu$ for all quanta absorbed by the atom in a sufficiently long time must be equal to the sum of the values of $h\nu$ for all quanta re-emitted by this atom. This condition, of course, follows from the law of conservation of energy.

Both the cases of non-coherence described above relate to non-coherence only *in frequency*, since the *continuity* of the wave process is not here destroyed.

(3) Non-coherence caused by pressure effects. This factor is apparently the most important one in stellar atmospheres, since, as we have seen in Chapters 12 and 13, the broadening of the levels of atoms in stellar atmospheres, with the possible exception of supergiants, is largely determined by pressure effects. A consideration of these effects shows that, in the case in question, the re-emission of the absorbed quanta must be completely *non-coherent in both frequency and phase*. This is most easily seen for the case of statistical broadening. If, for instance, the process of absorption of a quantum by an atom takes place at a moment when there is a perturbing particle close to the atom, the frequency of the absorbed quantum will differ from that of a quantum absorbed in the absence of the perturbing particle. At the moment when the electron returns to its original level, the perturbing particle may be so far from the atom that the frequency of the emitted quantum is different from that of the absorbed quantum. The difference (excess or defect) of energy between the absorbed and emitted quanta goes to increase (or decrease) the energy of the perturbing particle, whose kinetic energy (at infinity) will now differ from its original value by exactly the difference in energy of the absorbed and emitted quanta (if, of course, the absorbing atom returns to its original state and non-coherence in the *unperturbed* atom can be neglected).

In this case, the frequencies of the absorbed and emitted quanta are completely independent, and the re-emission of quanta is completely non-coherent. These results are applicable, for example, to the broadening of hydrogen lines by interatomic electric fields. The case of line broadening by collisions is similar to that of statistical broadening.

* The interlocking effect is also significant for subordinate lines (see above).

It follows from the above that, in the presence of pressure effects, the frequencies of the absorbed and emitted quanta must in general be different. *Scattering* processes are therefore *completely non-coherent* in the presence of pressure effects. Furthermore, the non-coherence here considered is non-coherence in both frequency and phase. In fact (see Chapter 11), pressure effects are caused by just the change in phase of the oscillations of the absorbing and emitting atoms.

Let us now consider how the equation of transfer should be modified in the presence of non-coherent scattering processes. Here we shall consider only non-coherence caused by the effects of collision damping. We have seen in Chapters 12 and 13 that in many cases these effects are the principal ones in the broadening of absorption lines.

In this case, the scattering is completely non-coherent. This means that the frequency distribution of the re-emitted quanta is independent of the frequency distribution of the radiation which excites the atoms. Consequently, the frequency distribution of the re-emitted quanta is determined by the law (11.23), i. e. the energy re-emitted in completely non-coherent scattering should be proportional to σ_ν, the **absorption coefficient in the line***.

Let us now consider an actual example of the construction of the equation of transfer for resonance lines. Here it must be taken into account that in general the broadening of these lines is determined both by pressure effects and by radiation damping. The rigorous separation of these two factors in writing the equation of transfer poses a problem as yet unsolved. However, if we start from the model of a classical oscillator, this separation can be performed as follows. Let the radiation damping constant for the given line be γ_{coh} and the collision damping constant γ_c. In this case the absorption of radiation is determined by the absorption coefficient (11.24) with $\delta_{ik} = (\gamma_{\text{coh}} + \gamma_c)/4\pi$, since $\gamma_1 \approx 0$ for resonance lines. Analysis then shows that, of the energy absorbed, the fraction $\gamma_{\text{coh}}/(\gamma_{\text{coh}}+\gamma_c)$ is *coherently* re-emitted, i. e. the frequencies of the absorbed and emitted quanta are equal, while the fraction $\gamma_c/(\gamma_{\text{coh}} + \gamma_c)$ is completely *non-coherently* re-emitted, where, in accordance with what was said above, the frequency distribution of the re-emitted energy is identical with the frequency dependence of s_ν (or σ_ν).

The corresponding emission coefficient j_ν consists of two parts. The first part, which corresponds to coherently scattered light, has the form

$$(j_\nu)_{\text{coh}} = \frac{\gamma_{\text{coh}}}{\gamma_{\text{coh}} + \gamma_c}\, \sigma_\nu \int I_\nu \frac{d\omega}{4\pi}\,. \tag{14.8}$$

---

* Since the probability distribution laws mentioned are the same for both emission and absorption processes.

The second part can be written by starting from the total amount of energy absorbed by one gram of matter, which is

$$4\pi \int \sigma_{\nu'}\, \mathrm{d}\nu' \int I_{\nu'} \frac{\mathrm{d}\omega}{4\pi} . \tag{14.9}$$

Of this energy, the fraction $\gamma_c/(\gamma_{\mathrm{coh}} + \gamma_c)$ is scattered non-coherently. Since we are concerned with *completely non-coherent scattering*, we can write for the second part of j_ν

$$(j_\nu)_{\mathrm{non}} = \frac{\gamma_c}{\gamma_{\mathrm{coh}} + \gamma_c} \sigma_{\nu,n} \int \sigma_{\nu'}\, \mathrm{d}\nu' \int I_{\nu'} \frac{\mathrm{d}\omega}{4\pi} , \tag{14.10}$$

where $\sigma_{\nu,n}$ is the absorption coefficient normalised to unity, i. e.

$$\sigma_{\nu,n} = \sigma_\nu \Big/ \int \sigma_\nu \, \mathrm{d}\nu . \tag{14.11}$$

(Here σ_ν is the absorption coefficient in the line, defined in the usual manner.) The integral of (14.10) over the whole line should give us the expression (14.9) multiplied by $\gamma_c/(\gamma_{\mathrm{coh}} + \gamma_c)$.

Consequently, if we take no account of true absorption processes, the equation of transfer, instead of (9.30), takes the form

$$\begin{aligned} \cos\theta \frac{\mathrm{d}I_\nu(\theta)}{\varrho\, \mathrm{d}h} &= (\varkappa_\nu + \sigma_\nu) I_\nu(\theta) - \\ - \frac{\sigma_\nu}{\gamma_{\mathrm{coh}} + \gamma_c} &\left\{ \gamma_{\mathrm{coh}} \int I_\nu \frac{\mathrm{d}\omega}{4\pi} + \frac{\gamma_c}{\int \sigma_\nu \mathrm{d}\nu} \int \sigma_{\nu'}\, \mathrm{d}\nu' \int I_{\nu'} \frac{\mathrm{d}\omega}{4\pi} \right\} - \varkappa_\nu B_\nu . \end{aligned} \tag{14.12}$$

The addition to the right-hand side of (14.12) of additional terms with the factor ε_ν cannot be made without a special analysis to find how far the recombinations to the kth level can alter the second term in the braces in (14.12).

The exact solution of the equation (14.12) for the case $\gamma_{\mathrm{coh}} = 0$, i. e. the solution of the equation

$$\cos\theta \frac{\mathrm{d}I_\nu(\theta)}{\varrho\, \mathrm{d}h} = (\varkappa_\nu + \sigma_\nu) I_\nu(\theta) - \frac{\sigma_\nu}{\int \sigma_\nu \mathrm{d}\nu} \int \sigma_{\nu'}\, \mathrm{d}\nu' \int J_{\nu'} \frac{\mathrm{d}\omega}{4\pi} - \varkappa_\nu B_\nu , \tag{14.13}$$

has been explicitly obtained by V. V. Sobolev [155]. This case is of great interest since, as we have seen in Chapters 12 and 13, in many cases we actually have $\gamma_{\mathrm{coh}} \ll \gamma_c$.

The absorption line contours constructed on the basis f this solution become more broadened as we approach the limb of the disc; this agrees with the results of observation. The calculations of G. Münch for the K line of Ca II, based on the solution of equation (14.12) with

certain assumptions, lead to a satisfactory agrement with the observed variation of the K line contour over the solar disc. The fact that the theory of line formation with non-coherent scattering is in better agreement with observation than the theory for coherent scattering follows also from more recent calculations by V. V. SOBOLEV [160]. Thus taking account of non-coherent scattering largely removes the discrepancies between the theoretical and observed contours, though much further work is of course necessary. Further, it is very important to develop the physical foundations of the theory of non-coherent scattering by the use of laboratory results*. It must be mentioned that the existing laboratory data completely confirm the hypothesis of the non-coherence of scattering processes in the presence of broadening collisions. W. ORTHMANN and P. PRINGSHEIM have shown (using mercury lamps) that, when the atoms of a gas are excited by the radiation of a narrow resonance line, the new line which arises from re-emission processes is considerably broader than the former one, and has a contour which is entirely determined by broadening processes due to the collisions of the atoms in question with other atoms.

We have now considered processes of non-coherent scattering related to pressure effects. If the non-coherence arises only from the broadening of the lower level (subordinate lines), analysis shows that the scattered radiation is not completely non-coherent, but consists of two parts. The first part corresponds to coherent scattering and the second to non-coherent scattering. However, subordinate lines also are, apparently, broadened mainly by pressure effects. Hence we should expect that the considerations given regarding completely non-coherent scattering would be applicable to these lines also.

In conclusion, the following remark must be made. Both observation and theory (though still incomplete) show that the difference between line contours in the cases of coherent and non-coherent scattering is relatively small. This means that all our quantitative results regarding such parameters as the turbulent velocity, the damping constant, the number of absorbing atoms (and thus the chemical composition), etc., ought not to be much altered by replacing coherent scattering processes by non-coherent ones. In other words, all our quantitative results are apparently of the right order of magnitude.

3. Central residual intensities in absorption lines. Comparison of theory and observation. Let us now pass to the subject of the residual intensities in the central parts of absorption lines. We have already said on several occasions that there is a serious divergence between the results of calculations and those of observation. Namely, for the centres ($\nu = \nu_0$)

* For the theory of the diffusion of resonance radiation, taking account of the frequency change in absorption and emission processes, see L. BIBERMAN [22].

of fairly strong lines in the solar spectrum, where, according to (10.20), we should expect (for $\varepsilon_\nu = 0$) the result $r_{\nu_0} \approx 0$, the observed values of r_{ν_0} lie in general between 0·05 and 0·2. Such discrepancies exist not only for the Sun, but also for other stars. On the basis of (10.20) we should always expect the result $r_{\nu_0} \approx 0$ to hold for fairly strong lines if $\varepsilon_\nu = 0$. It is easy to see this by substituting a sufficiently large value of η_{ν_0} in this formula. The extent of the disagreement between the results of the theory in question and observation for stars of non-solar type can be seen, for example, from Fig. 43, compiled for the H_γ line by G. A. SHAĬN [144] from Simeis spectrograms. On the axis of abscissae are placed the spectral types from O to M 0 and $\log_{10} T_e$, and on the axis of ordinates, the depth $R_{\nu_0} = 1 - r_{\nu_0}$. It may be seen from this graph that, between the classes F 5 and B 2, where the H_γ line must be considered fairly strong, the value of R_{ν_0} differs very appreciably from unity, and that of r_{ν_0} from zero.

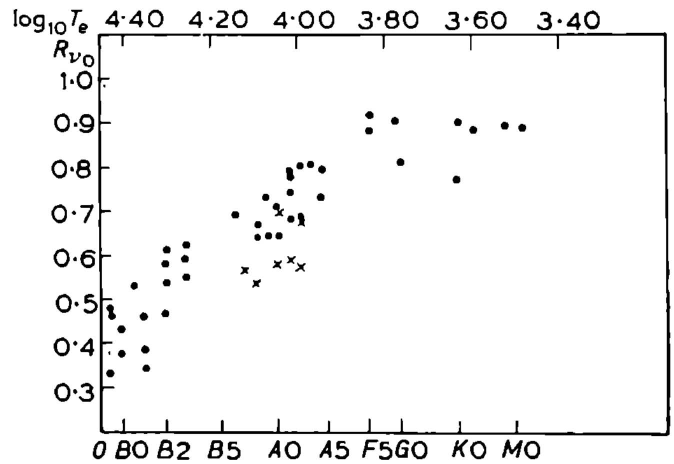

FIG. 43

Let us now consider a possible interpretation of these results. We first ascertain to what extent the intensity within a line is determined by processes of true absorption. We have seen in Chapter 9 that, when processes of true absorption are predominant ($\varepsilon_\nu \approx 1$), the intensity of radiation in the centre of a strong absorption line must be close to $B_\nu(T_0)$, for all angles θ. Thus, if the intensity of radiation in the parts of the continuous spectrum near the line is $B_\nu(T_E)$, where T_E is the temperature of the radiation (see Chapter 8), then to a sufficient approximation we can put

$$r_{\nu_0} \approx B_\nu(T_0)/B_\nu(T_E) \,. \tag{14.14}$$

(In many cases it can be assumed that $T_E \approx T_e$.)

Consequently, in this case r_{ν_0} will have a quite appreciable value, different from zero. It is true that in Chapter 9 we showed that the central parts of absorption lines in the solar spectrum are formed mainly not by processes of true absorption, but by those of scattering. However, some part ε_ν of the absorbed quanta *is* re-emitted according to the laws of true absorption. Moreover, it is evident that the introduction of a value $\varepsilon_\nu > 0$ in (10.20) increases r_{ν_0} in comparison with the case

$\varepsilon_\nu = 0$. We shall now consider this possibility. Here we shall as yet have in mind only lines of the solar spectrum; we shall speak later of stars of other spectral classes.

In obtaining the residual intensity r_{ν_0} in the centre of a strong line we shall work from the following simplified considerations. Since, for the line considered, the absorption coefficient σ_ν is very large in its central parts, we observe radiation coming from a very thin surface layer of the solar photosphere. Inside this layer, which includes part of the lower regions of the chromosphere, we shall assume the temperature to be constant (see Chapter 21) and equal to the surface temperature T_0. In view of the small thickness of the layer considered, we shall assume the values of $\eta_\nu = \sigma_\nu/\varkappa_\nu$ and of ε_ν to be independent of depth also. Consequently, we can use formula (10.18), putting $b_\nu = 0$ because of the assumed constancy of the temperature. For the flux H_ν^0 (0) which corresponds to the continuous spectrum, we can take

$$H_\nu^0(0) = B_\nu(T_{\mathrm{E}}) \,, \tag{14.15}$$

where T_{E} is the temperature of the radiation. The value of T_{E}, in the case where the linear expansion (8.10) is justified, can be found from (14.15) and (10.19). In general this can be done by equating $\pi B_\nu(T_{\mathrm{E}})$ to the theoretical flux πH_ν in the radiation of the continuous spectrum at the boundary of the atmosphere. Finally, T_{E} may also be determined from observation.

Taking this into account, we find for r_ν

$$r_\nu = \tfrac{4}{3}\sqrt{3}\,\frac{\sqrt{(1+\varepsilon_\nu\eta_\nu)}}{\sqrt{(1+\eta_\nu)}+\tfrac{2}{3}\sqrt{[3(1+\varepsilon_\nu\eta_\nu)]}}\,\frac{B_\nu(T_0)}{B(T_{\mathrm{E}})}\,. \tag{14.16}$$

For $\varepsilon_\nu = 1$ we obtain approximately* the equation (14.14), as we should expect.

It should also be noted that formula (14.16) cannot be applied for $\eta_\nu = 0$, since it has been obtained for the case $\eta_\nu \gg 1$, i. e. for the centres of strong lines. The use of (14.16) for $\eta_\nu = 0$ is permissible only if $T = T_0 =$ constant throughout the atmosphere of the star. In the centre of a fairly strong line, not only η_ν, but also $\varepsilon_\nu\,\eta_\nu$, is considerably greater than unity. In this case formula (14.16) takes the form

$$r_\nu \approx \tfrac{4}{3}\sqrt{3}\,\frac{B_\nu(T_0)}{B_\nu(T_{\mathrm{E}})}\sqrt{\varepsilon_\nu} = 2{\cdot}3\,\frac{B_\nu(T_0)}{B_\nu(T_{\mathrm{E}})}\sqrt{\varepsilon_\nu}\,, \tag{14.17}$$

where it is assumed that $\varepsilon_\nu \ll 1$.

* The factor 1.07 which then appears before $B_\nu(T_0)/B_\nu(T_{\mathrm{E}})$ arises from the mathematical simplifications of the method, particularly the use of formula (14.15).

In the solar spectrum for $\lambda = 4500$ Å, $T_E \approx 6200°$; for $\lambda = 6000$ Å, $T_E \approx 6000°$. With $T_e = 5710°$, we find from formula (4.25) that $T_0 \approx 4640°$. With these values, we obtain for $\lambda \approx 4500$ Å the ratio $B_\nu(T_0)/B_\nu(T_E) \approx 0{\cdot}17$, and for $\lambda = 6000$ Å, $B_\nu(T_0)/B_\nu(T_E) \approx 0{\cdot}30$. Thus, in the former case a factor close to 0·4, and in the latter case one close to 0·7, stands in front of $\sqrt{\varepsilon_\nu}$ in formula (14.17). It follows from this and from formula (14.17) that, if for example $r_{\nu_0} \approx 0{\cdot}05$ in the centre of an absorption line lying near $\lambda = 4500$ Å, then we should have $\varepsilon_\nu \approx 0{\cdot}016$, and for a line lying near $\lambda = 6000$ Å, $\varepsilon_\nu \approx 0{\cdot}005$.

Let us now consider what values of ε_ν we should expect for strong lines in the solar spectrum. For definiteness, we shall discuss resonance lines. We shall denote the lower and upper levels of such a line by the symbols 1 and k respectively.

We recall that the quantity ε_ν is the proportion of the total selectively absorbed energy that is transformed into thermal energy. It is clear that this quantity is also the proportion of the total number of transitions, accompanied by the absorption of quanta, which correspond to processes of true absorption. Since we wish to obtain only an approximate value of r_{ν_0}, we shall not take account of transitions upwards from the level k to higher discrete levels.

Let us now consider electrons which pass from the ground level to the level k in consequence of absorption processes (transitions of the type $1 \rightarrow k$). Having arrived at the level k, these electrons may leave it in three ways: (1) by transitions back to the ground level; the number of such transitions referred to one excited atom is equal to the coefficient A_{k1}; (2) by photo-ionisation processes from the level k; the number of such processes is denoted by C_{kf}, and it is equal to the expression (8.6) divided by $n_{r,k}$; (3) by collisions of the second kind, which transfer electrons from the kth level to the ground level without the emission of a quantum $h\nu_{1k}$. Let the number of such processes referred to one excited atom be c_{k1}.

Of these three processes, the first corresponds to scattering, and the latter two to the transformation of radiant energy into thermal energy. Since the total number of transitions from the kth level, referred to one atom, is $A_{k1} + C_{kf} + c_{k1}$, and the number of transitions corresponding to true absorption is $C_{kf} + c_{k1}$, we have

$$\varepsilon_\nu = (C_{kf} + c_{k1})/(C_{kf} + c_{k1} + A_{k1}) \,. \tag{14.18}$$

It is evident from the meaning of this formula that the value of ε_ν must be the same everywhere in the line*.

* See the article by B. Strömgren [164] for a more rigorous justification.

In formula (14.18) we can neglect the term c_{k1} in both the numerator and the denominator. For we have seen that the central parts of absorption lines are formed in the outermost layers of the solar atmosphere. In these layers the particle density is small, and consequently collisions of the second kind cannot be important there. This conclusion is confirmed by direct calculation. Thus the value of $\varepsilon_\nu = \varepsilon$ can be calculated from the formula

$$\varepsilon_\nu = \varepsilon = C_{kf}/(C_{kf} + A_{k1}) \,. \tag{14.19}$$

To find the quantity C_{kf} we can return to formula (8.6). If we use the approximate formula (10.16) at the boundary, then we have by (8.6)

$$C_{kf} \approx \frac{2\pi}{h} \int\limits_{\nu_k}^{\infty} (H_\nu)_{\tau_\nu=0} \frac{(k_\nu')_k}{\nu} \, d\nu \,. \tag{14.20}$$

Here we can put $(H_\nu)_{\tau_\nu=0} = B_\nu(T_E)$, where T_E is the temperature of the radiation, to estimate C_{kf} in the region which is most important for the photo-ionisation of the given atom from the kth level. For the levels of the hydrogen atom, and for levels of other atoms which have small binding energies $\chi_{r,k}$, we can use formula (5.24) in calculating $(k_\nu')_k$. For low-lying terms with large $\chi_{r,k}$, direct quantum-mechanical calculations or laboratory data must be used in each case.

In Table 7 we give the values of C_{kf} obtained from formula (14.20), using the equation $(H_\nu)_{\tau_\nu=0} = B_\nu(T_E)$ and formula (5.24). In this last formula we take $Z = 1$, $g' = 1$ and $n = k = 2$.

Table 7

$\chi_{r,k}$ (eV)	C_{kf} $T_E = 4900°$	C_{kf} $T_E = 6500°$
2	2×10^5	$7{\cdot}5\times10^5$
3	$1{\cdot}5\times10^4$	9×10^4
4	1×10^3	$1{\cdot}2\times10^4$
6	5	225
8	0·03	4·5

Although the values in the table are approximate, we can nevertheless estimate from them the order of magnitude of ε. Let us first consider the D_1 and D_2 lines of sodium, for which, according to observation, r_{ν_0} is approximately 0·05 to 0·06. For these lines $\chi_{0,k} = 3{\cdot}03$ eV. If we take as an upper limit $T_E = 6500°$, we find $C_{kf} \approx 9 \times 10^4$. On the other hand, $A_{k1} \approx 6 \times 10^7$ for these lines. We hence obtain for the lines in question $\varepsilon \approx 0{\cdot}0015$, and $r_{\nu_0} \approx 0{\cdot}03$.

For the Ca I line $\lambda = 4227$ Å, with an observed $r_{\nu_0} \approx 0{\cdot}03$, the transition coefficient $A_{21} \approx 1{\cdot}4 \times 10^8$, $\chi_{0,k} \approx 3{\cdot}2$ eV, $C_{kf} \approx 6 \times 10^4$, $\varepsilon \approx 4{\cdot}3 \times 10^{-4}$, and the theoretical value of $r_{\nu_0} \approx 0{\cdot}007$. Thus in both cases the theoretical value of r_{ν_0} is less than the observed value. The situation is still worse for the H and K lines of ionised calcium. For these $\chi_{1,k} = 8{\cdot}7$ eV, and consequently (from Table 7) $C_{kf} < 1$, while for both lines $A_{k1} \approx 1{\cdot}4 \times 10^8$ per sec. We find from this that in the case considered ε is equal to or less than 10^{-8} in order of magnitude, which is negligibly small for the purpose of explaining the observed value of r_{ν_0}, 0·07 to 0·08.

4. **Deviations from thermodynamic equilibrium in the Sun. Stars of other spectral classes.** Let us attempt to take account of the effect of possible deviations from local thermodynamic equilibrium on the last term on the right-hand side of equation (10.2). Since it gives the emission corresponding to processes of true absorption, these deviations can be represented by simply multiplying $\varepsilon_\nu\, \eta_\nu\, B_\nu$ by some quantity Q. If $Q > 1$, then the emission will be greater than when the laws of thermodynamic equilibrium are valid; if $Q < 1$, the emission will be less. In this case, equation (10.2) takes the form

$$\begin{aligned}\cos \theta \, \mathrm{d}I_\nu(\theta)/\mathrm{d}\tau_\nu = (1+\eta_\nu)\, I_\nu(\theta) \, - \\ - (1-\varepsilon_\nu)\, \eta_\nu\, J_\nu - B_\nu - Q\, \eta_\nu\, \varepsilon_\nu\, B_\nu\,.\end{aligned} \tag{14.21}$$

In order to write the quantity Q explicitly, we consider the physical meaning of the calculations we have just performed. The increase of r_{ν_0} due to the difference of ε_ν from zero is caused by *recombinations* to the kth level, i. e. by captures of electrons into this level. Let us consider processes of the types $1 \to k \to f$ and $f \to k \to 1$. In thermodynamic equilibrium, the two types of process balance each other. In the outermost layers of the atmosphere of the Sun (or star), the intensity of radiation in the line is small, being weakened in the ratio r_{ν_0}; hence transitions of the type $1 \to k$, and consequently $k \to f$, take place less often than in thermodynamic equilibrium. The diminution of the intensity in the line frequency has, however, no direct effect on processes of the type $f \to k \to 1$. Each process of the type $f \to k \to 1$ leads to the production of a quantum in the line frequency. The predominance of these processes must increase the central intensity of the absorption lines.

The case we have considered is essentially an instance of **fluorescence**, i. e. of a process in which quanta of high energy are transformed into quanta of low energy. In fact, in the case considered the number of transitions corresponding to the closed cycle $1 \to k \to f \to 1$ is, for the reasons given, less than the number corresponding to the cycle $1 \to f \to k \to 1$. In other words, the processes of transformation of

quanta with the higher energy $h\nu_{1f}$ into quanta with the lower energies $h\nu_{fk}$ and $h\nu_{1k}$ take place much more often than the converse processes with the successive absorption of two quanta $h\nu_{1k}$ and $h\nu_{kf}$ and the re-emission of one quantum $h\nu_{1f}$.

On the basis of these considerations, let us consider the problem of the value of Q. Since almost every electron which arrives at the level k falls directly back to the ground level (the coefficient C_{kf} is much smaller than A_{k1}), producing additional radiation in the line centre, the intensity of this latter radiation is proportional to the number of recombinations to the kth level. Furthermore, in the presence of equilibrium the number of recombinations which take place in 1 cm^3 in 1 second must be equal to the number of photo-ionisations in the same volume and time. Hence an increase in this number of photo-ionisations for any reason must lead to an increase in the number of captures C_{fk}, and so to an increase in the number of additional transitions A_{k1}. On the other hand, the number of photo-ionisations which take place in 1 cm^3 is mainly determined by transitions of the type $1 \rightarrow f$, since, as we have seen, there are few atoms in the kth excited state. Consequently, the number of transitions $k \rightarrow 1$ which produce additional radiation must be proportional to the coefficient C_{1f}. Bearing this in mind, we can write for Q

$$Q = C_{1f}/C_{1f,t}\,, \tag{14.22}$$

where $C_{1f,t}$ is the number of photo-ionisations corresponding to thermodynamic equilibrium, and C_{1f} is the same number in the case considered.

A more rigorous discussion of this problem leads to the expression [164]

$$Q = \frac{C_{1f}}{C_{1f,t}} \Big/ \frac{C_{kf}}{C_{kf,t}}\,. \tag{14.23}$$

The difference between (14.23) and (14.22) is noticeable only in cases where the upper term k lies fairly deep. The H and K lines of Ca II, for which $\chi_{1,k} = 8{\cdot}7$ eV, belong to this class.

If we now assume that, within the relatively thin layer in which the observed central residual intensity is formed, the value of Q is constant, and that the quantities η_ν, ε_ν and $B_\nu(T) \approx B_\nu(T_0)$ are also constant (see above), then we can find the solution of equation (14.21), using the method of Chapter 10; instead of equation (10.5) we obtain the equation

$$\tfrac{1}{4}\, \mathrm{d}H_\nu/\mathrm{d}\tau_\nu = (1 + \varepsilon_\nu \eta_\nu)\,(J_\nu - \mu_\nu B_\nu)\,, \tag{14.24}$$

where

$$\mu_\nu = (1 + \eta_\nu \varepsilon_\nu Q)/(1 + \eta_\nu \varepsilon_\nu)\,, \tag{14.25}$$

while equation (10.6) remains unchanged. Further, using equations (10.6) and (14.24) and the fact that $\mu_\nu B_\nu = \mu_\nu B_\nu(T_0) =$ constant, we obtain instead of (10.10) the equation

$$d^2(J_\nu - \mu_\nu B_\nu)/d\tau_\nu^2 = q_\nu^2 (J_\nu - \mu_\nu B_\nu) \tag{14.26}$$

and its solution

$$J_\nu = \mu_\nu B_\nu(T_0) + C_\nu e^{-q_\nu \tau_\nu}. \tag{14.27}$$

Finally, determining the constant C_ν as in Chapter 10, we find for the flux $H_\nu(0)$ in the centre of a strong line the relation

$$H_\nu(0) = \frac{4}{3} \frac{1 + \eta_\nu \varepsilon_\nu Q}{1 + \eta_\nu \varepsilon_\nu} \frac{q_\nu B_\nu(T_0)}{1 + \eta_\nu + \frac{2}{3} q_\nu}. \tag{14.28}$$

Introducing also the flux $H_\nu^0(0)$ by the relation (14.15), we obtain for r_ν

$$r_\nu = \frac{1 + \eta_\nu \varepsilon_\nu Q}{1 + \eta_\nu \varepsilon_\nu} \frac{4}{3} \sqrt{3} \frac{\sqrt{(1 + \varepsilon_\nu \eta_\nu)}}{\sqrt{(1 + \eta_\nu)} + \frac{2}{3} \sqrt{[3(1 + \varepsilon_\nu \eta_\nu)]}} \frac{B_\nu(T_0)}{B_\nu(T_E)}. \tag{14.29}$$

If we now take into account the fact that, for the centre of a strong line, not only $\eta_\nu \gg 1$, but also $\eta_\nu \varepsilon_\nu \gg 1$, the first factor in (14.29) becomes equal to Q, and instead of (14.17) we have the formula

$$r_\nu \approx 2{\cdot}3\, Q \frac{B_\nu(T_0)}{B_\nu(T_E)} \sqrt{\varepsilon_\nu}. \tag{14.30}$$

Thus formula (14.30) differs from (14.17) by the factor Q.

On returning to the calculations of r_{ν_0} which we have carried out, we see that, to explain the observed value of r_{ν_0} in the case of the D lines of Na I, we must have $Q \approx 2$, and for the line 4227 Å of Ca I, $Q \approx 4$. Thus, by introducing even relatively small values of Q, agreement can be achieved between theory and observation. To do so, it is only necessary to assume [see formula (14.22)] that the density of radiation which ionises the atoms of Na I and Ca I from the ground level is a few times greater than the density of radiation in thermodynamic equilibrium. However, it is impossible to explain in this way the central residual intensities of the H and K lines in the solar spectrum. The application of formula (14.17) here gives for r_{ν_0} a value of the order of 10^{-4} or less, i. e. the value of r_{ν_0} obtained from observation ($r_{\nu_0} \approx 0{\cdot}07$ to $0{\cdot}08$) is more than 700 times greater than the theoretical value of r_{ν_0}. Consequently, it would be necessary in this case to have Q not less

than 700, and this is very dubious. The greatest difficulty is that the value of $\chi_{1,k} = 8{\cdot}7$ eV for the H and K lines is little different from the ionisation potential of Ca II ($\chi_1 = 11{\cdot}8$ eV). Hence the two parentheses in (14.23) cannot differ by very much. Consequently, for lines where $\chi_{r,k}$ is large, the theory we have considered is apparently untenable. How then can we explain the observed values of r_ν at the centres of the H and K and similar lines? A new possibility has been pointed out recently. V. V. Sobolev [155, 160] has shown that processes of non-coherent scattering also lead to an appreciable increase of r_{ν_0}. The same result has been obtained by I. W. Busbridge [26]. It is possible that the anomalously high intensities in the centres of many absorption lines in the solar spectrum will be explained in just this manner. Moreover, processes of non-coherent scattering may perhaps explain the observed values of r_{ν_0} for the D_1 and D_2 lines of Na I, 4227 Å of Ca I, etc., without the introduction of the hypothetical factor* Q, though in general fluorescence processes (with $Q = 1$) should play a considerable part for some lines. It must be pointed out that these processes may occur not only on account of the decrease of the radiation density in some member of the cycle of transitions (in the case considered above, the transition $1 \to k$ was this member). As will be shown in Part IV, which deals with the planetary nebulae, fluorescence must occur in every case where the dilution factor W, which appears in formula (8.4), is less than unity. This is so, in particular, for the outer layers of the atmosphere of the Sun (and of stars), since there, as we know, $W \approx \frac{1}{2}$.

To study fluorescence phenomena, the equations of the steady state are constructed, each of which equates the total number of transitions in 1 cm^3 in 1 second, to any given level of the atom in question, to the number of transitions in 1 cm^3 in 1 second starting from this level. The solution of these equations as written for the different levels furnishes the answer to the problem proposed.

The efficiency of fluorescence processes should be increased in the presence of factors which increase the ionisation of atoms. These considerations should apparently be important for the H_α, H_β, ... lines of the solar spectrum (see Part IV). In carrying out a quantitative analysis, it must be recalled that in the majority of cases the effect of fluorescence is diminished by the selection rules.

In conclusion, we must make the following remark. We have seen that in almost all cases the quantity $\varepsilon_\nu = \varepsilon \ll 1$. This means, however, that on going no great distance from the line centre, we shall have not $\varepsilon_\nu\, \eta_\nu \gg 1$, but on the contrary the inequality $\varepsilon_\nu\, \eta_\nu \ll 1$. Conse-

* In fact, observations and theoretical calculations indicate that in the region of photo-ionisation of Na I and Ca I it would be correct to use the second column of Table 7 and consider Q as not greater than unity.

quently, by (10.20) and (10.9) we can neglect the effect of ε_ν for the transitional parts and wings of lines, and put $\varepsilon_\nu = 0$ there.

It must be mentioned that, although we can put $\varepsilon_\nu \approx 0$ in the wings, their whole formation in the case of strong lines corresponds rather to the case of thermodynamic equilibrium than to scattering. In fact, we shall see in Chapter 15 that for the centre of a strong line $J_{\nu_0} \approx B_\nu$ almost throughout the atmosphere, with the exception of its outermost parts, and this means (cf. Chapter 15) that the excitation of atoms takes place in accordance with the laws of thermodynamic equilibrium. The sum $(1-\varepsilon_\nu)\ \sigma_\nu J_\nu + \sigma_\nu \varepsilon_\nu B_\nu$ in equation (9.29) is equal to $\sigma_\nu B_\nu$ both in the case of simultaneous fulfilment of the equations $J_\nu = B_\nu$ and $\varepsilon_\nu = 0$, and in the case of fulfilment of the single equation $\varepsilon_\nu = 1$.

According to (10.20) and (10.9), we can often put $\varepsilon_\nu \approx 0$ for faint lines, where η_ν is small even in the centre of the line, if, of course, ε_ν is not close to unity, so that $\varepsilon_\nu\ \eta_\nu < 1$. We have already used this result in the preceding chapters. In general, however, caution must be exercised in any such case. In particular, for lines with a small binding energy $\chi_{r,k}$ and a small transition coefficient A_{ki}, the quantity ε_ν may, by (14.19) and Table 7, approach unity. This is true to an even greater degree in hot stars, where the density of ionising radiation is large, and consequently the coefficients C_{kf} are also large.

We shall now consider briefly the question of the central intensities of absorption lines in the spectra of stars of non-solar type. As we go from the Sun to hotter stars, the density of radiation in the far ultraviolet region of the spectrum rapidly increases, and the part played by fluorescence processes thereby increases also, especially in fairly hot stars with *extended envelopes*, where the dilution factor may satisfy even the inequality $W \ll 1$. In certain conditions, fluorescence leads to the appearance of bright lines. This case will be specially considered in Part VI.

The increase in the density of ultra-violet radiation as we pass to the hot stars brings about an increase in the value of C_{kf}, and thereby of ε_ν. Hence, in fairly hot stars, the value of ε_ν may become equal to unity even for strong lines where A_{k1} is large. In this case r_{ν_0} is determined by formula (14.14). Thus strong lines, of many different elements but lying in nearly the same spectral region, should give (for these stars) nearly the same value of r_{ν_0}. This is confirmed by observation [144]. The surface temperature T_0 can be estimated from the central intensities of these lines.

Finally, it is to be noted that an additional factor increasing r_{ν_0} might be the increase in the kinetic temperature in the outer layers of the atmospheres of some classes of stars [70, pp. 114-5].

Chapter 15. Methods of studying the chemical composition of stars. The results of investigations

1. The application of the elementary theory of curves of growth to the study of the chemical composition of stars. The mean electron concentration. The subject of the chemical composition of stellar atmospheres, and of the heavenly bodies in general, is of great cosmogonical significance. By comparing the compositions of various cosmic objects, we can draw conclusions concerning the evolution of the universe around us. For this reason, the question of the chemical composition of the heavenly bodies always attracts particular attention. The study of the chemical composition of stellar atmospheres is regarded as one of the principal tasks of the theory of stellar atmospheres.

It must be noticed that the finding of the chemical composition with fair accuracy has to be preceded by the establishment of some "initial" approximation to the chemical composition, since the physical state of the matter in the stellar atmosphere will itself depend on its chemical composition. It is extremely difficult to solve simultaneously the two problems, that of the structure of the stellar atmosphere, and that of finding the chemical composition of the atmosphere. Hence we shall use the method of successive approximations: each step towards the solution of the problem of the chemical composition of the atmosphere refines our knowledge of its physical structure and makes possible a further step towards finding the chemical composition.

At the present time, two paths have been marked out in the study of the chemical composition of stellar atmospheres. The first of these is based on the application of the elementary theory of curves of growth (Chapter 12), while the second is based on the study of line contours of different elements (Chapter 13). The second method may also include the construction of curves of growth, but on a more exact basis. Bearing in mind the inadequacy of the contemporary theory of curves of growth, mentioned in Chapter 12, we must regard the second method as considerably more exact than the first. However, the first method has hitherto, for a number of reasons, found greater application.

We have already given in Chapter 12 the general ideas relating to the determination of the chemical composition from curves of growth. Namely, the values of L in equation (12.37) are found from observation, and the numbers N_r from equation (12.38). In order to find the number of all atoms of the given element above 1 cm^2, it is necessary to take account of all stages of ionisation, i. e. to find the sum ΣN_r. This can be done as follows. If all the stages of ionisation of the atoms in question which are of any importance have absorption lines in the spectral region

considered, then the value of $\sum N_r$ is obtained by adding the values of N_r obtained directly from the curve of growth.

If, on the other hand, some ionisation state r, in which there is a considerable number of atoms, is not represented by observable absorption lines, the number N_r can be determined from N_{r-1} or N_{r+1} (whichever is known) by means of Saha's formula (5.11). To do this the temperature T and the electron pressure p_e must be known.

Here the values of T and p_e must be some *mean values* for the whole atmosphere, and this, of course, at once introduces some indeterminacy. In particular, it is not yet clear which temperature should be inserted in Saha's formula: the effective temperature T_e, or one specially determined. In one particular case it can be determined as follows. If we have found from the curve of growth the ratio of the numbers N_r for two elements in *two successive* ionisation states (e.g. N_r/N_{r-1} and N_{s+1}/N_s), then, writing Saha's formula once for each element, we have two equations in two unknowns, T and p_e. However, this method often leads to differing values of T and p_e, if we use not one but several pairs of elements. Hence the mean values of T and p_e have to be taken from the values determined from several pairs of elements, and this is not a satisfactory method; it only displays the inadequacy of the contemporary theory of curves of growth.

If we take some definite value, for instance T_e, for T in Saha's formula, then p can be determined from elements where the ratio of the numbers of atoms in two successive ionisation states is known. Here, too, different values of p_e are usually obtained for different elements (sometimes the difference amounts to one or two orders of magnitude), so that we have to go back to averaging. These results again display the inadequacy of the method in question.

Another method of finding p_e has recently come into use. It is known that, for instance, the coalescence and disappearance of the lines of the Balmer series in the spectra of stars takes place before the theoretical series limit (at $\lambda = 3646$ Å) is reached. For example, the transition from lines to continuum in the spectra of stars of class A 0 takes place at approximately $\lambda = 3700$ Å, so that the continuum begins not at $\lambda = 3646$ Å, but at $\lambda = 3700$ Å.

The reason for the premature disappearance of the series lines, before the limit is reached, is the line broadening by the *pressure effect*. The appropriate investigations have shown that the principal part here is played by the broadening of the higher levels of the atoms by *electric fields* (statistical broadening). We recall (see Table 6) that the broadening of the level increases with its principal quantum number, so that after some term in the series the lines must coalesce. This coalescence usually begins with those lines of the series for which the binding energy

of the corresponding levels is small. Hence the probability that an electron from these levels will return to its original level is practically zero. On the contrary, it is much more probable that the electrons will be removed from these levels, i.e. the atom will be ionised. The cause of the ionisation may be either of two: (1) ionisation by radiation, (2) ionisation by the interatomic electric fields which broaden the levels (see Chapter 8). In both cases, the processes of line formation in the series correspond to *true absorption*. On the other hand, the attenuation of the radiation in the frequencies of the continuous spectrum of stars is likewise due to processes of true absorption. Thus the region where the absorption lines coalesce is in every respect equivalent to an ordinary part of the continuous spectrum beyond the series limits.

Of the two causes of ionisation mentioned, the second is, as a rule, the principal one. Hence it is quite clear that the coalescence and disappearance of lines in the series, as a result of the pressure effect, should depend on the concentration of charged particles in the atmosphere. The more of these particles there are in 1 cm^3, the less will be the number of distinguishable lines in the given series. Hence the number of distinguishable lines in the series is a measure of the electron concentration n_e, and consequently of the electron pressure p_e.

Let m be the principal quantum number of the last distinguishable line in the series. Then n (the number of charged particles in 1 cm^3) is determined by the following theoretical formula:

$$\log_{10} n = 23{\cdot}26 - 7{\cdot}5 \log_{10} m, \tag{15.1}$$

The number n which appears in formula (15.1) refers only to those charged particles which cause the coalescence of the lines in the series. It is found that, for temperatures below $10^5/m$, the number n includes both ions and electrons, so that in (15.1) we must put $n = 2n_e$. At higher temperatures, ions alone must be taken into account, and since the number of ions is equal to the number of electrons, in this case we must put $n = n_e$ in (15.1). It is true that electrons do not play a negligible part here*. However, this method of determining n_e is in general approximate and gives only the order of magnitude. Hence errors of 20 to 50 % are not considered large.

Laboratory investigations confirm formula (15.1); they introduce a small correction which increases the first term on the right of (15.1) by 0·10.

Formula (15.1) is applicable not only to hydrogen, but also to the alkali metals. In the latter case the effective principal quantum number must be used.

* For $T > 10^5/m$, the presence of electrons is a factor leading to an increase in collision damping.

Formula (15.1), as we have said, is approximate and gives only the order of magnitude. There are a number of factors which (spuriously) increase the value of n_e. The chief of these are, firstly, the imperfect resolution of the spectrograph and, secondly, Doppler effects connected with thermal motions, turbulence, rotation, etc. On the other hand, since the absorption coefficient is large at the series limit, we obtain n_e for the more external layers of the stellar atmosphere. Hence the values of n_e found in the manner described are on this account too small. The method in question is in general more justified for emission lines.

A third method of determining mean values of p_e is based on the calculation of models of stellar photospheres according to the general methods explained in Chapter 7. It is true that this gives not the mean value of p_e but its dependence on depth. However, we can determine the required mean value of p_e from such data if we know approximately the effective level at which absorption lines are formed.

As an example, we give in Table 8 a list of determinations of p_e for stars of various classes and luminosities. For some stars, and especially for the Sun, several determinations exist. The statement of these independent determinations is important because it permits a comparison of values of p_e found by various methods.

Table 8

Star	Class	p_e (bars)	Star	Class	p_e (bars)
α Ori	cM 2	$5{\cdot}1\times10^{-7}$	70 OphA	dK 0	1·2
α Sco	cM 1	$1{\cdot}3\times10^{-6}$	Sun	dG 3	18, 46, 17, 31
α Boo	gK 0	0·004, 0·003	θ Cyg	F 5	12
π Cep	G 5	3·5	α CMi	dF 3	160, 86
α Aur	G 0	3·4, 0·05	α CMa	dA 2	950, 120
γ Cyg	cF 8	0·5	γ Gem	dA 2	120
δ Cep	cF 5	1·2, 2·0	τ Sco	dB 0	1200
α Per	cF 4	7·4, 4·2	ϕ' Ori	B 0	630
α Car	gF 0	10	10 Lac	O 8·5	630

The most important feature seen from Table 8 is that p_e decreases with T_e; this is, of course, related to the general decrease in the ionisation of atoms.

Let us now return to the subject of the chemical composition of the atmospheres of stars. Having determined by some method the electron pressure p_e, we can, as we said above, take into account all states of ionisation of the given element and obtain the value of $N = \sum N_r$ for the element considered. By comparing N for different elements, we find the relative chemical composition (content) of the solar or stellar atmosphere, i.e. the quantities a_s and α_s [see (5.1) to (5.4)].

In cases where we do not know the oscillator strength and the theoretical intensity determined by formula (12.39), we can use the "solar" intensities mentioned in Chapter 12. To do so, it is simplest to use formula (12.40), which we have already applied to determine $(T_{ex})_{st}$. Having found, as in Chapter 12, the difference $\Delta = (\log_{10} X_0)_{st} - (\log_{10} X_0)_{\odot}$ for various lines, and having determined the dependence of ε_i on Δ, we find the quantity

$$\Delta = \log_{10} \frac{(N_r)_{st}\,(u_r)_{\odot}}{(N_r)_{\odot}\,(u_r)_{st}} - \log_{10} \frac{(v_0)_{st}}{(v_0)_{\odot}}$$

at the point where $\varepsilon_i = 0$.

Having determined $(v_0)_{st}$ and $(v_0)_{\odot}$ from the curves of growth which we have constructed, we find the ratio $(N_r)_{st}/(N_r)_{\odot}$ from the $\Delta\,(\varepsilon_i = 0)$ just obtained. We can thus discover the chemical composition of the stellar atmosphere relative to that of the solar atmosphere. Of course, all stages of ionisation must again be taken into account, i.e. the sums $\sum N_r$ must be compared.

The method just explained can also be applied to compare the chemical composition of two stars of neighbouring spectral classes. The corresponding results have been derived at Pulkovo by O. A. MEL'NIKOV (see Section 12.5).

It must be stated that the determination of the numbers N need not be based on formulae (12.37) and (12.38). If we have an insufficient number of multiplets to construct graphs like Fig. 34, the number $N_{r,i}$ can be found *directly* from the curve of growth. The transition from $N_{r,i}$ to N_r is then made by the use of Boltzmann's formula (5.54).

It has to be noted that this method of studying the chemical composition clearly assumes the validity of Boltzmann's Law. The validity of this law for not very large ε_i (up to 5 eV) can be tested by means of a graph similar to Fig. 34. Here the temperature T_{ex} can be determined directly from observation, namely from the slope of the line in Fig. 34. For large ε_i, however, above 4 to 5 eV, the possibility of applying Boltzmann's Law to stellar atmospheres is by no means evident. Hence, in the case where the observed lines of the element considered correspond to transitions from very high energy levels ($\varepsilon_i > 5$ eV), the application of the method in question may lead to errors. Such lines include, in particular, those of H, He, C, N, O, Si, S, etc., especially the lines of the ionised elements. Finally, the absorption lines of helium are entirely absent in the spectra of low-temperature stars, including the Sun. All such cases must be considered individually. Thus, for example, in the atmospheres of late-type stars, including the Sun, the relative content of H, C, N, and O can be found from a study of the molecular bands of the compounds CN, C_2, OH, NH, and CH. Moreover, very weak for-

bidden lines of oxygen have recently been discovered in the spectrum of the Sun; they correspond to transitions from low energy levels of the oxygen atom. Consequently, the application of Boltzmann's formula in this case ought not to lead to large errors. We shall speak later of other methods of determining the hydrogen content. The helium content in the atmospheres of low-temperature stars is generally taken as equal to that in the solar atmosphere, where it is determined for the chromosphere and the prominences. Here it is supposed that the chemical composition of the chromosphere and prominences is the same as that of the layers of the solar photosphere lying beneath (see Chapters 20 and 21).

2. **The distribution of atoms among excitation states.** It follows from what has been said above that the hypothesis of Boltzmann's distribution plays a considerable part in the problem of the chemical composition of stellar atmospheres. It is therefore necessary to consider this hypothesis more closely. It can be tested, in particular, by means of such graphs as Fig. 34. The linearity of the relation between ε_i and $\log_{10}(N/g_i)$ indicates the applicability of Boltzmann's Law to stellar atmospheres. However, even so there are things which are still uncertain. For instance, the marked difference in the values found for T_{ex} and T_e, which we have discussed in Chapter 12, has not yet been explained.

What are the predictions of theory in this case? First, it is necessary to notice the following general characteristic of ionisation and excitation processes in stellar atmospheres: both these processes are determined almost entirely by the radiation field, and not by collisions between particles. (For excitation processes, this can easily be seen by means of calculations like those which we performed in Chapter 8 for ionisation processes.) The density of matter in stellar atmospheres is too small for excitation by collisions to be at all efficient. In such a case the main condition for the validity of Boltzmann's formula is the *identity* of the exciting radiation field with the radiation field in thermodynamic equilibrium. In other words, the quantity J_ν at a given point must be equal to Planck's function $B_\nu(T)$ at that point. This is a necessary and sufficient condition for the existence of a Boltzmann distribution, in the absence, of course, of external perturbing factors (for example, a strong electric field).

Let us consider whether this condition is fulfilled in stellar atmospheres. To do so, without going beyond general principles, we can use the model which we discussed in Section 10.2, i. e. the model with a linear expansion of the function B_ν. From (10.11), (10.17) and (8.10) we can write

$$J_\nu = B_\nu + \frac{\frac{2}{3} b_\nu - B_\nu(T_0)(1+\eta_\nu)}{\frac{2}{3}\sqrt{[3(1+\eta_\nu)(1+\varepsilon_\nu\eta_\nu)]} + 1 + \eta_\nu} e^{-\sqrt{[3(1+\eta_\nu)(1+\varepsilon_\nu\eta_\nu)]}\,\tau_\nu} . \quad (15.2)$$

The number of transitions $n_{i\to k}$ made by n_i atoms in 1 second under the action of the radiation field J_ν is, by (3.6),

$$n_{i\to k} = 4\pi n_i \int J_\nu \frac{\sigma_\nu}{h\nu} \,\mathrm{d}\nu \approx \frac{4\pi}{h\nu} n_i \int J_\nu \sigma_\nu \,\mathrm{d}\nu , \tag{15.3}$$

where the integration is to be extended over the whole of the line which corresponds to the transition $i \to k$.

Introducing the expression (15.2) into (15.3) and using formula (11.39) for the selective absorption coefficient, we can find the dependence of the product $J_\nu \sigma_\nu$ on the frequency within the part of the spectrum occupied by the line, and thereby ascertain which parts of the line are the most efficient in exciting the atom. Such calculations, performed for various ε_ν, $\varkappa_\nu$, etc., show that the *central* parts of the line are the most efficient, since the absorption coefficient has, according to (11.39), a *sharp* maximum at the centre of the line (the Doppler core). Thus, in studying the deviation of J_ν from B_ν, we need pay attention only to the most central parts of absorption lines.

To study the difference $J_\nu - B_\nu$, we consider a fairly strong line, in whose centre $\eta_{\nu_0} \gg 1$. Further, we at first assume (what is often found) that $\varepsilon_\nu \ll 1$, and $\varepsilon_\nu \eta_{\nu_0} \ll 1$. In this case, formula (15.2) for the line centre can be written (see Chapter 8 regarding b_ν):

$$J_{\nu_0} \approx B_\nu (T) - B_\nu (T_0) \exp [- \sqrt{(3\eta_{\nu_0})}\, \tau_\nu] . \tag{15.4}$$

It follows from (15.4) that at the boundary of the atmosphere, where $\tau_\nu = 0$ and $T = T_0$, the quantity $J_{\nu_0} \approx 0$. This result, which we have already encountered in a previous chapter, indicates that the theoretical central residual intensity of strong lines with $\varepsilon_\nu \ll 1$ is almost equal to zero. Thus, at the boundary of the stellar atmosphere, the exciting radiation does not correspond to radiation in thermodynamic equilibrium. However, as soon as we move into the atmosphere, even to $\tau_\nu \approx 0{\cdot}1$, the exponential in (15.4) becomes considerably less than unity. If, for example, $\eta_{\nu_0} = 10^4$, then for $\tau_\nu = 0{\cdot}1$ we have $\sqrt{(3\eta_{\nu_0})}\, \tau_\nu \approx 17$, and $e^{-17} \approx 4 \times 10^{-8}$. Even for $\tau_\nu = 0{\cdot}02$, the value of $\exp[- \sqrt{(3\eta_{\nu_0})}\, \tau_\nu] \approx 0{\cdot}03$. In other words, we should expect the equation $J_{\nu_0} \approx B_\nu(T)$, where T is the local temperature, to hold practically throughout the stellar atmosphere. If we no longer assume that $\varepsilon_\nu \ll 1$ and $\varepsilon_\nu \eta_{\nu_0} \ll 1$, the approach to the equilibrium state is made even more rapidly, since when $\varepsilon_\nu \to 1$ the conditions approximate more and more closely to those of thermodynamic equilibrium. In fact, for $\varepsilon_\nu = 1$, but again with $\eta_{\nu_0} \gg 1$, we have instead of (15.4):

$$J_{\nu_0} \approx B_\nu (T) - \frac{B_\nu(T_0)}{1 + \frac{2}{3}\sqrt{3}} \exp [- \sqrt{3\eta_{\nu_0}}\, \tau_\nu] . \tag{15.5}$$

In particular, at the boundary of the atmosphere, where $\tau_\nu = 0$ and $T = T_0$, we have $J_{\nu_0} \approx \frac{1}{2} B_\nu(T_0)$. The factor $\frac{1}{2}$ is equivalent to the dilution factor at the boundary*. The appearance of η_{ν_0} before τ_ν in the exponent (instead of $\sqrt{\eta_{\nu_0}}$) makes the transition to thermodynamic equilibrium even more rapid. Thus, for strong lines, and to a somewhat lesser degree for medium lines, we may expect Boltzmann's formula to be applicable practically throughout the stellar atmosphere (for the temperature corresponding to each level, of course).

For faint lines $\eta_{\nu_0} < 1$ or $\eta_{\nu_0} \ll 1$, i. e. the radiation field approaches what it would be in the absence of absorption lines. It is evident that under these conditions the deviation of J_{ν_0} from $B_\nu(T)$ will be determined chiefly by the presence of the gradient $\mathrm{d}B_\nu/\mathrm{d}T$.

Observations give us an idea of the *mean* excitation state over the *whole* atmosphere. The linear relation between $\log_{10}(N_i/g_i)$ and ε_i which is usually obtained (for $\varepsilon_i < 5$ eV) indicates that in such cases the excitation conditions are not far from those of thermodynamic equilibrium. This fact is certainly related to the comparative smallness of the gradient $\mathrm{d}B_\nu/\mathrm{d}\tau_\nu$, for the observable region of the spectrum, in the atmospheres of ordinary stars and of the Sun; this follows from theoretical calculations also. For large ε_i, however, great caution must be exercised, since these ε_i correspond to the far ultra-violet regions of the spectrum, where the gradient $\mathrm{d}B_\nu/\mathrm{d}\tau_\nu$ is very large in some cases. Finally, we must not forget the as yet unexplained fact that, according to observation, $T_{\mathrm{ex}} < T_e$, i. e. the mean excitation temperature is found to be less than the effective temperature. In particular, calculations show that the application of formula (15.2) cannot explain this inequality. Either the approximate equality $T_{\mathrm{ex}} \approx T_e$ or the inequality $T_{\mathrm{ex}} > T_e$ is usually obtained.

Observations usually give graphs like that shown in Fig. 34. However, in a number of cases there are deviations from linearity. These are observed chiefly for cool red giants of the α Bootis type. It is found that in the atmospheres of these stars the number of atoms in high quantum states is considerably greater than would be the case for conditions of thermodynamic equilibrium. Thus the Boltzmann distribution does not exist here. Some lines of Fe I in the infra-red part of the Sun's spectrum, for example lines whose lower term is $e^5\mathrm{D}$, also point to a deviation from the Boltzmann distribution. The nature of all these deviations is as yet unknown.

3. Refined methods of studying the chemical composition of the stars. Faint lines. The method of studying the chemical composition of stellar atmospheres which is based on the application of the elementary theory

* The slight difference of the expression $\frac{2}{3}\sqrt{3}$ from 1 arises from the inexactitude of the methods used to derive (15.2).

of curves of growth has a number of defects, due largely to the schematic nature of the theory itself (see Chapter 12). The principal one is that a very imperfect model of the scattering "reversing layer" above the "photosphere" is used in this theory, the thickness of the layer being taken the same for all lines. In reality, this "thickness", and consequently the value found for the "number of absorbing atoms above the photosphere", varies with wavelength on account of the variation of the continuous absorption coefficient with wavelength. Moreover, direct analysis shows that the "thickness of the reversing layer" is different for lines with different excitation and ionisation potentials.

A second method of studying the chemical composition is based on a model which takes account of scattering and of true absorption of radiation (both general and selective). In this case we start from the equations of transfer (9.29), (9.32), (10.2). The general principles of this second method of studying the chemical composition are as follows. Having taken some provisional chemical composition, obtained, say, by using the first method, we find p, p_e, ϱ, T, $\varkappa_\nu$, τ_ν, etc. for every point of the stellar photosphere, using for stars the method of Chapter 7, and for the Sun that of Chapter 17. Then, using the general method which we explained at the beginning of Chapter 13, we form an expression for the quantity η_ν. This quantity should be a function [see (13.18)] of the content a_s of the element in question and of the remaining parameters p, p_e, ϱ, T, etc.

Now, taking several values of a_s close to the expected value (we suppose a_s constant throughout the stellar atmosphere), we integrate (see Chapter 13) the corresponding equation of transfer and construct the contour of the line under consideration for each of the a_s taken. By comparing these theoretical contours with that of the given line, obtained from observation, we find the a_s required. Here a comparison can be made both of the contours themselves and of the equivalent widths. To do so, it is necessary to calculate first the equivalent widths W_λ for the whole assembly of theoretical contours. Having determined the content (i. e. the value of a_s) for every element, we can compute the structure of the photosphere afresh, with greater exactness.

For a more accurate estimate of the value of a_s, the application of the *method of curves of growth* is recommended. This can be done, for instance, as follows. Let us consider the group of lines which begin from some one lower level i, and let the oscillator strengths f_{ik} be known for all these lines. Then, giving a_s some value close to the expected value, we can construct a contour and calculate the total absorption for each of the lines considered, in accordance with the considerations just given. We thereby obtain the theoretical relation between f_{ik} and W_λ. Now, placing $\log_{10} f_{ik}$ on the axis of abscissae and $\log_{10}(W_\lambda/\lambda)$ for

the corresponding lines on the axis of ordinates, we construct a portion of the theoretical curve of growth. We plot on the same graph the relation between the logarithms of the same values of f_{ik} and W_λ/λ for the corresponding lines, as obtained directly from observation. If the value taken for a_s were the correct one, the two curves of growth thus constructed would coincide. However, this will not in general happen. The theoretical curve is somewhat displaced (along the axis of abscissae) relative to the curve obtained from observation. From the best coincidence of the two curves we evidently determine the correction $\Delta \log_{10} a_s$. Adding this to the value of $\log_{10} a_s$ taken previously, we find an improved value of a_s.

The following remarks must be made regarding the above.

(1) Since the value of a_s gives the content of the element in question relative to hydrogen, we can, having found a_s for all elements whose content is fairly large, estimate the percentage content of hydrogen also.

(2) In cases where T_e and g are unknown, we must first determine them in order to calculate the structure of the photosphere. We can determine T_e by using the scale of effective temperatures (or by means of the methods explained in Chapter 16), and g either as in Chapter 13 (comparison of theoretical and observed contours of H_γ and H_δ), or by using the mass, determined in some way (e. g. from the mass-luminosity relation), and the radius. (This is most reliably done for eclipsing variables where the spectra of the two components can be distinguished.)

The second method of studying the chemical composition considers two cases which must be distinguished: firstly, strong and medium lines, and secondly, faint and very faint lines. Analysis of the two cases shows that the use of faint and very faint lines to solve the problem of chemical composition often has great advantages over the use of strong and medium lines. These advantages are chiefly the following. (1) Contrary to what is the case for strong lines, the intensity of faint lines is independent of the damping coefficients γ_c, which have not yet been accurately determined, and which vary with depth in the stellar atmosphere. (2) The central parts of strong lines are distorted, much more than those of faint lines, by various factors which we discussed in the last chapter (fluorescence, non-coherence) and which are difficult to take into account. (3) In the construction of a fairly accurate contour of a strong or medium line, we usually encounter a number of difficulties. For faint and very faint lines, on the other hand, we can use the solution (10.43), which is more exact than (10.20). (4) Only a very small number of chemical elements have strong lines in the region of the spectrum of the Sun and stars which is accessible to investigation. The use of lines of medium intensity is inconvenient, since these lines lie on the flat part of the curve

of growth, where the equivalent width of the lines varies very little with the number of absorbing atoms. Hence even a small error in the equivalent width found may give a very large error in the values of N_r or $N_{r,i}$. Because of the circumstances just enumerated, faint lines should, as we have said, be preferred to strong ones, though in many cases, particularly in the analysis of spectra with small dispersion, the strong and medium lines are an important means of chemical analysis of stellar atmospheres.

The general method of finding a_s is practically the same for both faint and strong lines. However, let us consider how faint lines may be used [88], having in mind the Sun, for which the possibilities of studying the chemical composition are greatest, since apparatus with a large dispersion can be employed.

We shall start from the solution (10.43). For the Sun, the relation between B_ν and τ_ν can be determined from the law of darkening to the limb of the disc, found from observation. The weight functions $G_1(\tau_\nu)$ and $G_2(\tau_\nu)$ can also be determined in this way. Calculations show that the difference between these functions is not large. In the red region of the spectrum, the two functions practically coincide. These properties of the functions also follow from an analysis of the expressions (10.44) and (10.45), if we take into account the fact that $E_1(x) > E_2(x)$ everywhere. Hence we can take instead of $G_1(\tau_\nu)$ and $G_2(\tau_\nu)$ in (10.43) their mean, which we denote by $\bar{G}(\tau_\nu)$. By so doing we introduce into the final result errors which do not exceed 20 % on the average, and these are quite admissible with the problem of the chemical composition of stellar atmospheres in its present state. Thus, instead of (10.43), we can write

$$R_\nu \approx \int_0^\infty \frac{\varkappa_\nu{}^0 + \sigma_\nu{}^0}{\varkappa_\nu} \bar{G}(\tau_\nu)\,\mathrm{d}\tau_\nu = \int_0^\infty \frac{\sigma_\nu}{\varkappa_\nu} \bar{G}(\tau_\nu)\,\mathrm{d}\tau_\nu\,; \tag{15.6}$$

the last member of this equation follows from (9.31).

We introduce, in place of $\sigma_\nu/\varkappa_\nu$ in (15.6), the expression (13.18) for it. Bearing in mind that in the photosphere of the Sun (at least up to 4000 Å) the continuous absorption is determined mainly by H^- ions, we can neglect the second term in the denominator of (13.18). As a result we obtain

$$R_\nu = a_s \int_0^\infty \frac{s_\nu Z_{r,i}(p_e, T)}{k_\nu(\mathrm{H}^-)\,p_e} \bar{G}(\tau_\nu)\,\mathrm{d}\tau_\nu\,. \tag{15.7}$$

Integrating (15.7) over frequency and using (12.12), we find:

$$\frac{W_\lambda}{\lambda} = \frac{W_\nu}{\nu} = \frac{\int_0^\infty R_\nu\,\mathrm{d}\nu}{\nu} \approx \frac{a_s}{\nu_{ik}} \int_0^\infty \frac{\bar{G}(\tau_\nu)\,Z_{r,i}(p_e, T)}{k_\nu(\mathrm{H}^-)\,p_e} \left\{ \int_0^\infty s_\nu\,\mathrm{d}\nu \right\} \mathrm{d}\tau_\nu\,, \tag{15.8}$$

where the integration over frequency takes place within the limits of a very narrow (faint) absorption line.

In place of the integral in the braces in (15.8) we can use (11.5), replacing B_{ik} by f_{ik} with the aid of (11.14). As a result we obtain, instead of (15.8),

$$\frac{W_\lambda}{\lambda} \approx a_s \frac{\pi e^2 f_{ik}}{c m_e \nu_{ik}} \int\limits_0^\infty \frac{\bar{G}(\tau_\nu)\, Z_{r,i}(p_e, T)}{k_\nu(\mathrm{H}^-)\, p_e}\, d\tau_\nu \,. \tag{15.9}$$

The dependence of $\bar{G}(\tau_\nu)$, and thus of T, on τ_ν is derived from observations, which give the dependence of $B_\nu(T)$ on τ_ν. That of p_e on τ_ν can be found from the theory of photospheres. In this way the integral on the right-hand side of (15.9) can be evaluated by methods of numerical integration. Consequently, having determined W_λ from observation and knowing f_{ik}, we obtain the required value of a_s directly. Of course, in this method also it is necessary to assume some provisional chemical composition of the solar atmosphere, since the value of p_e is determined by the ionisation of both hydrogen and the metals, the approximate content of which has to be known beforehand. However, if a_s is determined from lines of neutral metals, this can be avoided, as we shall now see.

We multiply together the ionisation equation (5.11) and the equation (5.54) with $r = 0$. Then, using (5.6), we find

$$\frac{n_{0,i}}{n_1} = p_e \frac{g_{0,i}}{u_1} \frac{h^3}{2(2\pi m_e)^{3/2}} \frac{1}{(kT)^{5/2}}\, e^{\chi_{0,i}/kT} \,, \tag{15.10}$$

where $\chi_{0,i}$ is the binding energy of level i. The majority of the metal atoms in the solar atmosphere are *singly* ionised, so that $n_0 \ll n_1$. Hence we can suppose that the *total* number n of atoms of some element in 1 cm^3 is approximately equal to the number of *ionised* atoms of that element in 1 cm^3. Consequently, we can assume that $n = n_0 + n_1 + n_2 + \ldots \approx n_1$. For $Z_{0,i}(p_e, T)$ we can therefore write the equation

$$Z_{0,i}(p_e, T) = n_{0,i}/n \approx n_{0,i}/n_1 \,; \tag{15.11}$$

using (15.10), we find for $Z_{0,i}(p_e, T)$ the expression

$$Z_{0,i}(p_e, T) \approx p_e \frac{g_{0,i}}{u_1} \frac{h^3}{2(2\pi m_e)^{3/2}} \frac{1}{(kT)^{5/2}}\, e^{\chi_{0,i}/kT} \,. \tag{15.12}$$

Now, introducing this function into (15.9), we see that p_e cancels, so that the expression under the integral sign is a function of *temperature alone*. On the other hand, the relation between the temperature and τ_ν is found directly from observation. Thus, the integral on the right-hand side of (15.9) can be evaluated from observation alone. Consequently,

faint lines of neutral metals give directly the content a_s of the metal concerned relative to hydrogen.

4. **The results of investigations.** By means of the general methods just explained, with various modifications of course, a large amount of work has been done on the chemical composition of the Sun and stars. Various methods have been used for the Sun, and this makes it possible to estimate the reliability and accuracy of the results obtained. A comparison has shown that in a number of cases there are fairly large discrepancies in the values of a_s. These discrepancies are caused chiefly by differences in the values taken for the excitation temperatures and the oscillator strengths f_{ik}. However, because of the abundance of material, we now have sufficiently reliable data on the values of a_s for the Sun and for some stars. The probable error in the values of $\log_{10} a_s$ for these objects amounts, on the average, to 0·3 to 0·5, although in particular cases it may be considerably greater. This is especially true of elements where the lines studied begin from very high levels, i. e. from levels with large $\varepsilon_{r,i}$. These elements, as we have said, include He, C, N, O, Si, S, etc.

The values of a_s for the Sun, given in Table 1 (see Section 5.1) were obtained (except for He, C, and N) by W. J. Claas [33], who used mainly the method of faint lines. The helium content in the Sun is put equal to the helium content in the chromosphere and prominences (see Chapters 20 and 21). The carbon and nitrogen contents are based on the study of molecular bands.

To obtain the numbers N_s of atoms of various elements above 1 cm^2 of the solar photosphere, it is sufficient to know the number N_{II}, since $N_s = a_s N_{\mathrm{II}}$. Different authors give somewhat different values for N_{II}, lying between 10^{24} and 5×10^{24}. For rough calculations we can take $N_{\mathrm{II}} = 2{\cdot}5 \times 10^{24}$ atoms above 1 cm^2. It must be recalled that the quantities N_s give the *total* number of atoms above 1 cm^2, taking into account *all* degress of ionisation and excitation.

As well as for the Sun, the determination of the chemical composition has been carried out for stars of many different classes. The most reliable determinations are those for F-type main-sequence stars and giants, hot B-type stars like τ Scorpii (see Table 1) and O-type stars. The relative chemical composition (values of a_s) of the atmospheres of all these bodies, including the Sun, is approximately the same*. Furthermore. an examination of Table 1 shows that this chemical composition is similar to that of the planetary nebulae and of the interstellar gas. This, of course, is a fact of great cosmogonical importance. It cannot be ignored by any theory of cosmogony which professes to explain the possible ways of stellar evolution.

* It is not yet possible to speak of complete identity, because of the insufficient accuracy of the existing data.

It must be admitted that the question whether all stars of the main sequence have identical chemical compositions cannot yet be considered as finally resolved, since there are a number of difficulties. For instance, the position as regards helium is very indefinite. According to Table 1, the content of helium relative to hydrogen is about 0·2. However, according to recent investigations (see, for example, L. NEVEN and C. DE JAGER [112]), the helium content in the atmospheres of B-type stars is only 0·05. It is still difficult to say whether there is an actual difference in helium content between the atmospheres of B stars and that of the Sun, since the helium content of the Sun's atmosphere is itself not reliably determined. R. MICHARD [87] finds a lower content of helium in the Sun than was previously supposed.

Let us now consider the question of marked deviations from the "normal" composition. Here we must be extremely cautious, since there are a number of factors which distort the intensities of absorption lines and may lead to incorrect values of a_s. These include: (1) anomalous excitation and ionisation conditions. The possibility that such anomalous conditions exist is indicated, for example, by the study of the exterior envelopes of the Sun (chromosphere, corona and prominences); (2) turbulent motions in stellar atmospheres. These motions have different effects on strong and on faint lines, and an inadequate consideration of them may lead to incorrect values of a_s; (3) such factors as the magnetic field may also change the intensities of lines. As an example we may quote the star HD 125248, where the observed magnetic field varies with time, in phase with the variations in intensity of lines of ionised europium. The cause of these synchronised oscillations in two quantities which are apparently independent lies in the fact that in a strong magnetic field the lines of europium are split into numerous Zeeman components; the blending of these components produces a broad, relatively intense line. Thus, by taking account of the general magnetic field, we can apparently explain quite a number of observed anomalies, especially in connection with rare elements.

We shall enumerate the cases where the analysis of absorption lines (and the general nature of the spectrum) leads to the conclusion that the chemical composition is "anomalous".

(1) The bifurcation of the spectral sequence at class K 2 into the oxygen branch (classes K and M) and the carbon branch (class N). The spectra of oxygen stars of class M are characterised principally by bands of titanium oxide TiO. Stars of class N are characterised by bands due to carbon compounds, in particular Swan bands (C_2) and those of cyanogen CN and carbon-hydrogen CH. Here we are concerned with an obvious difference in chemical composition: in the atmospheres of M-type stars there is much more oxygen than in those of N-type

stars, while in the atmospheres of N-type stars there is considerably more carbon than in those of M-type stars.

In their extensive investigations on the spectra of carbon stars, G. A. SHAĬN and V. F. GAZE [43] have found other very interesting properties of these stars. In particular, the D_1 and D_2 lines in the spectra of some carbon stars are approximately ten times as intense as in the spectra of M-type stars. This difference is apparently due to a difference in temperature, and perhaps to an actual difference in the sodium content. Judging by the intensities of the atomic lines of other metals, such as Fe, Cr, Mn, etc., their contents are about the same in stars of classes N and M.

(2) The division of Wolf-Rayet stars into the carbon branch and the nitrogen branch. We shall discuss points (1) and (2) again in the next chapter.

(3) "Metallic-line" stars. These stars form a large group of objects with contradictory spectral features. Thus, if we use the hydrogen lines or the K line of Ca II, the star can be placed in the spectral class A, but if we judge by the intensities of the majority of the lines of neutral and ionised metals, the star's spectrum corresponds approximately to class F 5. The analysis of the spectra of such stars by J. L. GREENSTEIN [50] shows that the content a_s of the elements Ca, Sc, Ti, Zr, V and Mg is anomalously low; for some of the stars it is ten times less than in the atmospheres of ordinary stars.

(4) A number of peculiar stars are now known, in whose spectra the lines of hydrogen are much fainter than in the spectra of ordinary stars of the same temperature, and in some cases do not appear at all. All these objects seem to be characterised by a fairly high luminosity.

Among the stars of the early classes O and B, the stars HD 160641 and HD 124448 are of this type. The hydrogen lines do not appear in their spectra. The stars v Sagittarii and HD 30353 are similar, but are of class A. Finally, the group of stars with anomalously weak hydrogen lines includes a few of lower temperature, whose spectra show molecular bands of C_2; they include variable stars of the R Coronae Borealis type, and other stars which have approximately the same spectrum but are not variable.

There is reason to suppose that the hydrogen content in the atmospheres of all these stars is in fact low, since it is difficult to imagine any physical mechanism which would explain the almost total absence of the Balmer lines (and of the Balmer discontinuity), the lines of the other elements being of normal strength.

It may be thought, and there are weighty reasons for doing so, that there exist real and considerable deviations from the "normal" chemical composition in the stars mentioned under (1), (2) and (4). The position

regarding the "metallic-line" stars is more indefinite. All the metals which show anomalously low contents (Ca, Sc, Ti, Zr, V and Mg) have very similar properties. In fact, the atoms of all these metals have two outer electrons. Moreover, quantum-mechanical calculations performed for Ca II have shown that its absorption coefficient extends anomalously far into the ultra-violet region of the spectrum, and hence, even for a small excess of radiation in the far ultra-violet region of the star's spectrum, the atoms of Ca II will be anomalously strongly ionised. Because of the similarity between the atoms of the elements mentioned and those of calcium, it may be assumed that, if there is an excess of radiation in the ultra-violet, there will be an anomalously strong second ionisation for these elements as well, and this leads to an apparent decrease in the values of a_s. Other elements, such as Fe, Mn, which also have two outer electrons, possess relatively large second ionisation potentials, so that here the ultra-violet excess will play a smaller part.

GREENSTEIN [50] has put forward another mechanism for the ionisation of the atoms Ca II, Sc II, etc. by protons, involving a resonance mechanism with "charge transfer": $H^+ + Ca^+ \rightleftarrows H + Ca^{++}$.

Both these mechanisms, however, require an anomalously high ionisation of hydrogen in the atmospheres of the "metallic-line" stars, which ought to be noticeable in the contours of the Balmer lines, and especially in the H_α line. An investigation of this line in the spectra of nine "metallic-line" stars by É. R. MUSTEL' and L. S. GALKIN [109] has shown that there is no such anomaly in the ionisation of hydrogen. It is therefore very probable that in the "metallic-line" stars also there is an actual deviation from the "standard" chemical composition.

The carbon N-type stars are extremely interesting as regards chemical composition. Thanks to the investigations of Academician G. A. SHAĬN, who was awarded a Stalin Prize in 1950, and the investigations which he conducted together with V. F. GAZE, we now know a very great deal about the spectra of these stars.

In 1940 G. A. SHAĬN [138] showed that the carbon isotopes C^{12} and C^{13} are present in the atmospheres of N-type stars. This was discovered by a very careful analysis of the molecular bands in the spectra of these stars. It is extremely difficult to detect isotopes by means of atomic lines. In the case of molecules, however, the displacement between the bands of two isotopes reaches quite perceptible values (up to several ångströms). Working on these principles, G. A. SHAĬN detected a number of new Swan bands in the spectra of N-type stars: $C^{12}C^{12}$, $C^{13}C^{12}$, $C^{13}C^{13}$. Subsequently, bands of CN corresponding to the isotopes C^{12} and C^{13} were discovered also.

A second extremely important fact established by G. A. SHAĬN is the anomalously high content of the isotope C^{13} relative to C^{12}. This

content varies, from star to star, between the limits 0·05 and 0·3 or 0·4. Thus, in the latter case there are two or three times as many C^{12} atoms as there are C^{13} atoms, while on the Earth the ratio of the numbers of C^{13} and C^{12} atoms is about 0·01. Consequently the isotope C^{13} is extremely rare on the Earth. The Sun, also, has a very small relative content $C^{13}:C^{12}$, it being certainly less than $^1/_{30}$. The question of the content of the isotopes C^{12} and C^{13} in the atmospheres of stars has a direct bearing on the problem of the sources of stellar energy. In the wellknown carbon cycle, the isotopes C^{12} and C^{13} play a very important part (see Part VII).

New and interesting data have recently been obtained concerning the presence in the Sun's atmosphere of such elements as lithium and beryllium, as well as rather more uncertain data on the presence of technetium and deuterium.

These data apparently indicate that there is no mixing of matter between the innermost layers of the Sun, where the temperature is of the order of ten million degrees, and its atmosphere. The calculations of J. L. Greenstein and E. Tandberg-Hanssen [51] show that, if there were a circulation of gases in the Sun embracing a region where the temperature exceeds 3·6 million degrees, the beryllium atoms would be rapidly disintegrated. If, however, we also take into account the anomalously low lithium concentration that has been found by observation, extremely narrow limits must be placed on the mixing time of the gases in the Sun and on the boundary of the mixed region, and this is very improbable.

This leads to the hypothesis that there are some processes as yet unknown which generate lithium and beryllium in the outer layers of the Sun. The whole subject is discussed in the Liége symposium on nuclear processes in the Sun and stars [80].

To conclude this chapter, we must mention the importance of studying the chemical composition of groups of stars having various kinematic and spatial characteristics. This is very important for the study of the evolution of stars in the Galaxy.

The comparison of the chemical composition of stars belonging to various sub-systems with appreciably different velocity dispersions is of very great interest. We recall that the velocity dispersion is related to the spatial characteristics of the sub-system. In general, the "flatter" sub-systems have a small velocity dispersion, and the more "spherical" sub-systems a large dispersion (see the review article by P. P. Parenago [119]). The difference in the spatial and kinematic characteristics of different sub-systems gives reason to suppose that the origin of stars in different sub-systems is different, so that we should expect the chemical composition of the stars in question to be different also. The data at present available confirm this.

The study of giants of classes G and K shows that, in the spectra of stars having a relatively large velocity dispersion, the CN bands are fainter than the same bands in the spectra of stars with a smaller velocity dispersion. The opposite effect (stronger bands in the spectra of stars with a large velocity dispersion) is observed for the CH bands. For dwarfs of class F, the ratio of the equivalent widths of lines belonging to CH and Fe I in the spectra of stars having a relatively large velocity dispersion is also increased. These facts (and several others also) can be explained on the assumption that the percentage content of the metals (and, to a lesser extent, of the metalloids O, C, N) in the atmospheres of stars with a large velocity dispersion is less than that in the atmospheres of stars with a smaller velocity dispersion.

Similar results are obtained when the atmospheres of the *subdwarfs* are compared with those of main-sequence stars. The difference in the spectra of the stars of these two groups can be explained as being due to a smaller content of calcium and iron (and perhaps of other metals) in the atmospheres of the subdwarfs, i. e. of stars belonging to a spherical sub-system. The percentage content of hydrogen in the atmospheres of stars having a large velocity dispersion is apparently, according to the above investigations, almost the same as in the atmospheres of stars with a small velocity dispersion.

Chapter 16. The interpretation of the spectral sequence. The absolute-magnitude effect. The scale of effective temperatures

1. **Introductory remarks.** We have seen in Chapter 15 that the chemical composition of the atmospheres of the overwhelming majority of stars, of very different classes, is approximately the same. Consequently, the differences between the spectra of stars within the spectral sequence are certainly not due to a difference in the chemical composition of their atmospheres. In fact, as soon as it appeared (in 1920), the theory of the ionisation of stellar atmospheres showed that the chief factor which causes the change in the character of the spectrum along the spectral sequence is a change in *temperature*. From this standpoint, the nature of the spectrum (i. e. the presence of given absorption lines, their relative intensity, etc.) of any star within the spectral sequence should be determined mainly by the temperature alone. These ideas are in accordance with the linear Harvard classification of spectra, which depends on only *one* parameter and is of a purely empirical nature.

From the physical point of view, the results just mentioned are quite evident, since, for a given chemical composition, the nature of the spectrum of a star is determined by the ionisation and excitation conditions in its atmosphere, and these depend mainly on the temperature of the star.

However, the word "mainly" must here be used with the following reservation. In reality, the ionisation, by formula (5.11), depends not only on the temperature, but also on the pressure p_e, though of course to a considerably smaller extent. Hence the spectra of two stars at the same temperature, but with different luminosities, and consequently with different average p_e, will be somewhat different. In general, as we shall see later, the spectra of a dwarf and a giant of the same spectral class are distinguished by a number of features. Besides differences in the degree of ionisation, there are also differences caused by the pressure effect (the atmosphere of the giant is more rarefied than that of the dwarf). For this reason, a more exact empirical classification of spectra would have to be at least two-dimensional — *two* parameters must be given in order to characterise the spectrum of a star. The spectral class may be taken as one of these; it is very sensitive to temperature variations. The other parameter may be any characteristic (or several such) which is very sensitive to variations of absolute magnitude* (even if not altogether free from the influence of the temperature). The first problem for the observer is to assign values of these two parameters x and y for each star. The second problem is to examine the nature of the dependence of these two empirical parameters on T and g:

$$x = f_1(T, g),$$

$$y = f_2(T, g),$$

and to find the values of T and g from the given x and y. To conclude this short introduction, we must recall that the linearity of the Harvard classification ceases abruptly at the class K2, since a ramification of the spectral sequence into three branches takes place: the K-M stars, the K-R-N stars and the S stars.

The spectra of the K-M stars are characterised by titanium oxide bands, which are faint in K5 stars but reach great intensity in the class M8. The spectra of K-R-N stars are characterised by carbon and cyanogen bands, which become stronger from R0 to R5, N0 and N3. No star is known whose spectrum contains both titanium bands and carbon bands at the same time. The spectra of S-type stars are characterised by

* For the main sequence, the absolute magnitude of a star at a given temperature is in a one-to-one relation (by the mass-luminosity law) with the acceleration g due to gravity.

zirconium oxide bands. As has been said in Chapter 15, there is here apparently an actual difference in chemical composition.

We must now, starting from the basic ideas of the theory of stellar atmospheres, investigate how the nature of the spectrum will change when the temperature and the acceleration due to gravity at the surface of a star vary. To solve this problem, a special branch of theoretical astrophysics was built up during the twenties of the present century—*the theory of ionisation of stellar atmospheres.* This theory, because of the relative imperfection of our knowledge, at that time, of the formation of absorption lines, developed for a long time almost independently of the theory of absorption-line intensities. However, such a separation, which is in essence completely artificial, cannot be maintained at the present day. In fact, in order to study the nature of a stellar spectrum for various T_e and g and a given chemical composition, we have to construct the contours of absorption lines for these T_e and g (see Chapter 13). Then we must consider the change in the intensities of these lines in the spectra of stars with different T_e and g, and compare the theoretical results with those of observation. Thus, in principle, everything reduces to calculations based on the contents of Chapter 13. It is true that such calculations will, of course, be very schematic, since they do not take into account such factors as turbulence in the atmospheres of stars, the rotation of a star, small differences in the chemical composition, etc. However, these problems already lead us into the field of a two-dimensional classification.

It follows from the above that, to interpret the most important characteristics of the spectral sequence, we can here restrict ourselves to a qualitative discussion of the problem, since quantitative calculations would be based on the methods of Chapter 13, which we have already considered. The formulae and methods of the old "theory of ionisation of stellar atmospheres" have lost almost all significance nowadays*.

2. **The variation of line intensities with increasing temperature.** Let us now investigate how the intensities of various absorption lines will vary, for a fixed value of g, as we pass from stars with a low T_e to those with a higher T_e. (The effect of varying g will be considered in the next section.)

The starting-point for what follows is that the intensity of any line in the spectrum of a star is the greater, the more atoms there are which produce this line (see Chapter 12). Hence, in order to study the qualitative change in the intensity of any line in stellar spectra along the spectral sequence, it is sufficient for us to consider the change in the number of these atoms.

* This theory, in its final form, began from the idea of the so-called effective optical depth, which was supposed to be the same for all absorption lines. However, it has recently been found that this depth varies for different lines from 0.2 to 2, depending on the excitation potential, the region of the spectrum, etc.

The most characteristic properties of the spectra of stars with the lowest effective temperatures are as follows:

(1) the presence of bands due to molecular compounds;

(2) the presence of absorption lines which start from the ground level, for example the line 4227 Å of Ca I. At low temperatures the number of atoms in excited states is small, by formula (5.54);

(3) lines of ionised elements will be relatively faint in the spectra of the stars in question.

Let us now ascertain how the nature of the spectrum changes as we pass from these cool stars to stars with higher T_e. The increase in temperature causes a decrease in the number of molecules, because dissociation processes are intensified, and consequently the molecular bands become fainter. Here the first bands to disappear will be those which are due to easily dissociated compounds (with a small dissociation energy). The same will happen for the atomic lines which start from the ground level of the neutral atom. The number of neutral atoms will decrease (slowly at first, and then more rapidly) as the ionisation increases.

The lines of neutral atoms which correspond to transitions from excited levels will behave differently. As the temperature increases, the number of such atoms will at first increase [formula (5.54)]. The intensity of the corresponding lines will therefore increase also. However, this increase will not continue indefinitely. As the temperature increases further, the total number of neutral atoms of the element considered begins to decrease, because of the continuously increasing ionisation. For this reason, the intensities of lines which correspond to transitions from excited levels of neutral atoms must have maxima for some temperature. The line then becomes fainter and disappears at some spectral sub-class. (The lines of the Balmer series form an exception; they are observed even in the spectra of O-type stars. This is due to the exceptionally high percentage content of hydrogen in stellar atmospheres.) It is easy to see that lines which correspond to transitions from the ground level should, as a rule, disappear later (i.e. at higher temperatures) than those which correspond to transitions from excited levels.

The position of the maximum intensity of the line must depend on the ionisation potential of the element in question and on the excitation potential of the level from which the transition begins which forms the line. The greater these two potentials, the higher will be the temperature at which the line reaches its maximum. Thus, for example, the maximum intensity of the lines of He I is observed in the spectra of stars of the B sub-classes (the lines of neutral helium first appear in the spectra of stars of the sub-classes A 1 and A 2), while the majority of lines of neutral atoms of metals which are relatively easily ionised and excited have their maxima in the classes M, K and G. This, in particular, explains why

metal lines* are predominant in the spectra of stars of these classes. The lines of elements which are not easily excited are either absent or extremely faint in these spectra. Hydrogen forms an exception: even its first excitation potential is very high (for the second level, where the Balmer series begins, $\varepsilon_{0,2} = 10{\cdot}16$ eV). The Balmer lines, however, are fairly intense even in the spectra of G, K and M stars; this is due to the extremely high relative content of hydrogen (see Chapter 15).

The presence of an intensity maximum for the lines of neutral elements must relate, as we have just seen, to excited levels. However, in certain cases lines produced by transitions from the first level may also have maxima. For, if as a result of the rise in temperature the atoms in question are released by the disintegration of molecular compounds which contain them, the intensity of the corresponding atomic lines should likewise increase. The subsequent ionisation of the atoms causes a maximum of the intensity and then a decrease.

Let us now consider lines due to ions. We first examine lines produced by transitions from the ground state of the ionised atom. As we pass to the atmospheres of hotter and hotter stars, an *increase in the number of ionised atoms* will take place (at the expense of the number of neutral atoms). Because of this, the intensity of ionic lines increases, and new lines appear whose intensity also increases. However, this increase in intensity cannot continue indefinitely. As the temperature rises further, the increase in the second ionisation of the atoms considered leads to the result that the number of singly ionised atoms, having reached a maximum, begins to decrease. Thus the intensities of lines produced by transitions from the ground state of ionised atoms must also have some maximum in the spectral sequence. This maximum is reached at a point where the lines belonging to neutral atoms of the given element are already very faint. Furthermore, the maximum must be relatively flat, since the next stage of ionisation usually requires a considerably higher temperature.

It is quite evident that those lines of ionised atoms which correspond to transitions from excited levels must also have a maximum, though of course at a higher temperature, since additional energy has to be expended on the excitation of the atoms. These maxima should be more sharply defined than those of lines which start from the ground levels of the ions.

For ionised atoms, the position of the maximum must also depend on the values of the ionisation and excitation potentials. The larger these are, the higher is the temperature at which the corresponding lines and their intensity maxima are observed.

* Not only of neutral metals, but also of ionised metals with small ionisation potentials.

The concepts which we have explained are wholly applicable to higher ionisation states also.

Thus, as we move along the spectral sequence from the coolest stars (among those known to us) to the hottest, we should observe a continuous replacement of some lines (and maxima) by others which have greater ionisation and excitation potentials. A general change takes place correspondingly in the appearance of the spectra of stars, from spectra crossed by a multitude of absorption lines (and at first by bands) with strong blending, to the spectra of early classes, which have relatively few lines. In fact, as we pass from cool stars to hotter ones, we are concerned on the whole with atoms which are more highly ionised or difficult to ionise (for example, He I). In atoms with large $\chi_0, \chi_1, \chi_2, \ldots$, however, the majority of the strong lines are situated in the far ultra-violet, which is a region of stellar spectra not yet accessible to us. Thus, for instance, the lines of doubly and triply ionised atoms are, in general, encountered relatively seldom in the accessible region of the spectrum. The relative "scarcity" of lines in the spectra of early classes is hereby explained.

It is interesting to note that it was previously supposed (until the twenties of the present century) that the disappearance of metal lines and the appearance of lines of the inert gases in the spectra of hot stars was due to the destruction of the metal atoms and the formation of inert gas atoms. In actual fact, everything is explained by the multiple ionisation of the metal atoms; for such ions, practically all the lines are in the ultra-violet region of the spectrum, which is inaccessible to us. The observed lines of the inert gas atoms start from highly excited levels, so that a fairly high temperature is required to produce a sufficient number of absorbing atoms.

All the conclusions given here, i.e. the presence of maxima and the general nature of the variation of spectra as T_e increases, are in complete accordance with observation. The earlier "theory of ionisation of stellar atmospheres" achieved some success as regards quantitative conclusions, principally in the construction of a scale of "ionisation" temperatures (see below). However, this theory was fundamentally too elementary and was based on data for the absorption coefficients which are now known to be incorrect. Calculations performed in the last few years, based on a direct construction of absorption-line contours, can be used with justification only for stars of the early spectral classes, where the absorption is determined by hydrogen. For the cooler stars of classes F, G, K and M, where the continuous absorption is produced by negative hydrogen ions, and later by metals and by molecules, there are no calculations which give the variation of the intensity with T_e. Earlier calculations were based on the supposition that the absorption in the atmos-

pheres of these stars is produced by metals and by hydrogen. Nor are there sufficiently well-founded calculations for very hot stars, where the absorption of radiation is determined by hydrogen, by helium, and by free electrons.

The study of the behaviour of various lines along the spectral sequence is important, and not only for the solution of the general problem of interpreting the spectral sequence. This interpretation is applicable in practice to a number of situations. One of these is the establishment of a scale of stellar temperatures. Let us assume that we are considering stars with almost the same value of the acceleration g due to gravity, but with various T_e. Having constructed the theoretical contours and computed W_λ from them for some line, we find the value of T_e for which the line concerned has the maximum value of W_λ. Let this be $T_e = (T_e)_m$. Next, having selected stars of various classes, but with approximately the same value of g, we determine spectrophotometrically the spectral class in which the value of W_λ for the given line is a maximum. Let this be, say, the class F5. We can consequently suppose that the effective temperature of class F5 is $(T_e)_m$. A similar determination of T_e can be made for many other lines also, and these will in general have a maximum W_λ for other spectral classes. We thus obtain what is called the scale of "ionisation" temperatures, i.e. the relation between the spectral class and the temperature of a star: the temperatures found in the manner described must be essentially the effective temperatures, as follows from the method by which they were found.

We must extend the scale of effective temperatures to groups of stars with various values of g, since, as we shall see later, the effect of this parameter on the values found for T_e may be considerable in some cases.

The establishment of the scale of "ionisation" temperatures was one of the main tasks of the old "theory of ionisation" of stellar atmospheres. However, because of the much simplified character of this theory, the T_e determined from different lines for one and the same spectral class was often different.

At present, the method of maxima has retained some significance only for hot stars of classes B and O, where a number of difficulties are encountered in the establishment of the temperature scale (see below). Of course, here too the old "theory of ionisation" has to be entirely replaced by calculations based on the integration of the general equations of transfer (9.29), (9.32) and (10.2), taking into account the main ideas of Chapter 13. We shall discuss the results of such calculations a little later.

3. **The effect of the acceleration due to gravity.** Having considered in outline the effect of a star's temperature on its spectrum, let us now turn to a study of the part played by the acceleration g due to gravity.

A difference in the value of g means, as we know, a difference in the electron pressure: the smaller g, the smaller p_e, and conversely. On the other hand, the degree of ionisation, according to formula (5.11), is greater, the smaller p_e. Consequently, if we have two stars with the same effective temperature, one being a dwarf and the other a giant, the ionisation of the atoms in the atmosphere of the giant will be greater than in that of the dwarf, i.e. the spectrum of the giant will be of an earlier class than that of the dwarf.

Let us consider, for instance, two groups of stars with different values of g, say g' and g'', where $g' > g''$. Then, as we move along the spectral sequence towards higher temperatures, all phenomena in the spectrum (disappearance of lines, maxima of lines, appearance of lines of ionised elements, etc.) which depend on the ionisation will take place at lower T_e in the second group of stars than in the first. These facts are easily illustrated by means of formulae (5.11) and (5.54). For, by using these formulae, we can calculate the ratio of the number of ionised atoms of some element to the total number of atoms of that element (in the same volume). If we are interested in r times ionised atoms in the kth excitation state, then this ratio, which we denote by $y_{r,k}$, is

$$y_{r,k} = \frac{n_{r,k}}{n} = \frac{n_{r,k}}{n_r} \cdot \frac{n_r}{n}, \tag{16.1}$$

where

$$n = n_0 + n_1 + n_2 + \ldots = \sum_r n_r, \tag{16.2}$$

the sum being taken over all ionisation states.

For the first factor on the right-hand side of (16.1), we have from formula (5.54)

$$\frac{n_{r,k}}{n_r} = \frac{g_{r,k}}{u_r} e^{-\epsilon_{r,k}/kT}. \tag{16.3}$$

The second factor must be calculated by means of the ionisation equation (5.11). Writing this equation for $r = 0$, $r = 1$, $r = 2$, etc., we obtain, as a result of appropriate multiplication of the equations,

$$\frac{n_1}{n_0} p_e = K_0, \quad \frac{n_2}{n_0} p_e^2 = K_0 K_1, \quad \frac{n_3}{n_0} p_e^3 = K_0 K_1 K_2, \text{ etc.}, \tag{16.4}$$

where the notation (5.17) is used. In accordance with (16.4), the expression (16.2) takes the form

$$n = S(T, p_e)\, n_0, \tag{16.5}$$

where

$$S(T, p_e) = 1 + \frac{K_0}{p_e} + \frac{K_0 K_1}{p_e^2} + \frac{K_0 K_1 K_2}{p_e^3} + \ldots. \tag{16.6}$$

Using (16.1), (16.3), (16.4) and (16.5), we obtain the following expressions for $y_{0,k}$, $y_{1,k}$, ..., $y_{r,k}$, etc.:

$$y_{0,k} = \frac{g_{0,k}}{u_0} \frac{e^{-\varepsilon_{0,k}/kT}}{S(T, p_e)}, \tag{16.7}$$

$$y_{1,k} = \frac{g_{1,k}}{u_1} \frac{K_0 e^{-\varepsilon_{1,k}/kT}}{S(T, p_e) p_e}, \tag{16.8}$$

$$y_{r,k} = \frac{g_{r,k}}{u_r} \frac{K_0 K_1 \ldots K_{r-1} e^{-\varepsilon_{r,k}/kT}}{S(T, p_e) p_e^r}, \tag{16.9}$$

etc.

As has been mentioned in Chapter 5, the levels of consecutive ionisation states are usually so far apart that, for a given p_e, only r and $r + 1$ times ionised atoms exist in an appreciable amount, i. e. for example, n_0 and n_1, or n_1 and n_2, etc.* Thus the sum (16.6) is determined mainly by two terms only. Hence we have in this case, for neutral atoms,

$$y_{0,k} \approx \frac{g_{0,k}}{u_0} \frac{e^{-\varepsilon_{0,k}/kT}}{1 + K_0/p_e}. \tag{16.10}$$

If we are dealing with neutral atoms in the ground state, then (16.10) takes the form ($\varepsilon_{0,1} = 0$)

$$y_{0,1} \approx \frac{g_{0,1}}{u_0} \cdot \frac{1}{1 + K_0/p_e}, \tag{16.11}$$

where we must remember that the partition function u_0 depends on the temperature. The same is true of u_1, u_2, etc. In some cases the variation of u_r with T may cause some displacement in the position of the maximum of $y_{r,k}$. This can be shown, for example, for lines which correspond to transitions from the ground level of the singly ionised atom. We can observe these lines only in cases where they lie in the accessible region of the spectrum, i. e. when the nearest excited level has a fairly small excitation energy. The second ionisation potential of the atom, however, is comparatively high. Hence, when the temperature increases, the ions first pass into the excited state, and only later into the second ionisation state. The decrease in the number of singly ionised atoms in the ground state, after the maximum, is caused in these cases not so much by second ionisation as by the excitation of atoms to the nearest level.

* It follows from an analysis of absorption lines in stellar spectra that atoms in three successive ionisation states may sometimes exist in comparable amounts in stellar atmospheres. In this case we are dealing not with the ionisation state in a volume element, but with the total effect produced by the different layers of the atmosphere.

In general, the part played by the individual terms in the sum (16.6) can be ascertained by calculating them directly, for a given p_e of course.

The application of formulae (16.7) and (16.8) to the problem in which we are interested, that of the effect of p_e on the variation of the quantities $y_{r,k}$, is shown in Fig. 44. On the axis of abscissae is placed the temperature, and on the axis of ordinates the ratio of $n_{r,k}$ to n, expressed as a percentage. The curves A and B show the dependence

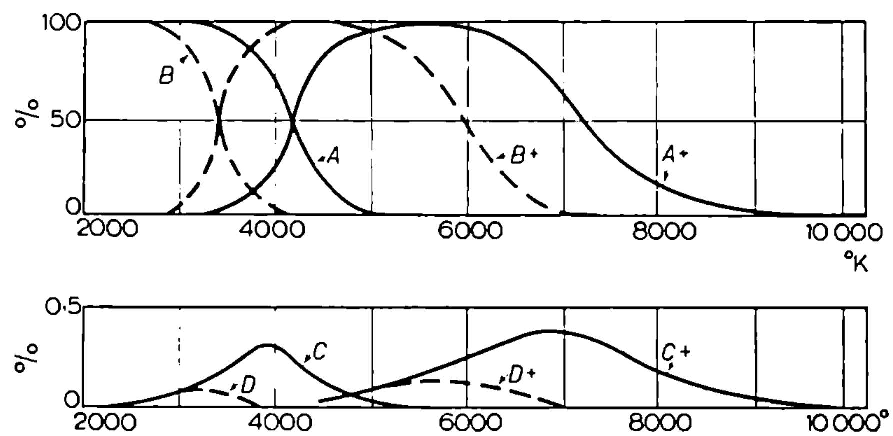

FIG. 44

of y on the temperature for the ground level of neutral calcium, from which starts, in particular, the transition which produces the resonance line 4227 Å. The curve A is constructed for $p_e = 10$ bars from formula (16.11), and the curve B for $p_e = 0{\cdot}1$ bar. The curves A_+ and B_+ show the dependence of y on the temperature for the ground level of ionised calcium, from which the transitions start which produce the H and K resonance lines. The curve A_+ is constructed for $p_e = 10$ bars, and the curve B_+ for $p_e = 0{\cdot}1$ bar. The curves A_+ and B_+ are constructed from formula (16.8), where $\varepsilon_{1,k} = 0$, since the level concerned is the ground level.

The curve C is constructed for $p_e = 10$ bars, and shows the dependence of y on the temperature for the excited level of neutral calcium from which the transition starts which produces the line 6162 Å. The potential of this level is about 1·9 eV. The curve D is calculated for $p_e = 0{\cdot}1$ bar. Both curves are constructed from formula (16.10) with $\varepsilon_{0,k} \approx 1{\cdot}9$ eV. Finally, the curve C_+ shows the dependence of y on the temperature for the excited level of ionised calcium with $\varepsilon_{1,k} \approx 3{\cdot}2$ eV, calculated for $p_e = 10$ bars. The transition which produces the line 3737 Å starts from this level. The curve D_+ is a similar curve for $p_e = 0{\cdot}1$ bar. Both curves are constructed from formula (16.8) with $\varepsilon_{1,k} \approx 3{\cdot}2$ eV.

The curves in Fig. 44 entirely confirm the statement made above that, as we move along the spectral sequence (in the direction of T

increasing), all the phenomena which depend on the ionisation take place at lower values of T in stars where g, and consequently p_e, is smaller, than in stars where g is larger.

Consequently, if we have two stars, a giant and a dwarf, of the *same* spectral class, the giant should have a lower effective temperature than the dwarf. For, if the spectral class is the same, this means, roughly speaking, that the *ionisation state of the atoms is the same*, while a lower temperature is needed to reach a given ionisation state of the atoms in the atmosphere of the giant, where p_e is less, than in that of the dwarf. This is very clearly noticeable for stars cooler than those of class F 5, i. e. for stars where the division into dwarfs and giants is clearly marked. We shall give in Table 11 quantitative data on the differences in T_e.

4. **The differences between the spectra of dwarfs and giants. The absolute-magnitude effect.** We have just said that a difference in T_e may be balanced by a difference in p_e, so that a giant and a dwarf may have practically the same spectral class. However, this compensation is not really complete. The similarity in the spectra of giants and dwarfs is only, so to speak, one of principle. There are a large number of marked differences of detail. These are due to the following causes.

(1) Let us write the ionisation formula (5.11) in the logarithmic form

$$\log_{10}\frac{n_{r+1}}{n_r} = \log_{10}\frac{u_{r+1}}{u_r} + \tfrac{5}{2}\log_{10}T - \frac{5040}{T}\chi_r - \log_{10}p_e + \log_{10}\frac{2\,(2\,\pi\, m_e)^{3/2}k^{5/2}}{h^3} \tag{16.12}$$

where χ_r is expressed in electron-volts. It follows from (16.12) that an increase or decrease in p_e respectively decreases or increases n_{r+1}/n_r by the same amount for all elements. On the other hand, a change in T causes a change in n_{r+1}/n_r which is not the same for different elements. For large χ_r, the change in n_{r+1}/n_r, for a given change in T, will be greater than for small χ_r. Thus the temperature "compensation" (decrease in T) in passing from a dwarf star to a giant star (of the same class) is too small for easily ionised atoms and too large for atoms which are difficult to ionise (χ_r is a factor of the most important term, $5040\,\chi_r/T$). Hence we should expect that, as we pass from dwarfs to giants, atoms with small χ_r will be ionised more strongly in the atmospheres of giants than in the atmospheres of dwarfs, and those with large χ_r less strongly; for elements with moderate values of χ_r the difference should be unimportant.

This should appear from the behaviour of absorption lines, and such a thing is in general observed. Thus, for instance, the lines of Ca I and Sr I are stronger in dwarfs and weaker in giants, while the reverse is

observed for the lines of Ca II and Sr II. (For Ca I $\chi_0 = 6{\cdot}09$ eV, for Sr I $\chi_0 = 5{\cdot}67$ eV, for Ca II $\chi_1 = 11{\cdot}82$ eV, and for Sr II $\chi_1 = 10{\cdot}98$ eV.)

(2) We know (see Chapter 13) that the contour of any absorption line, and also its equivalent width, are determined by the ratio η_ν:

$$\eta_\nu = \sigma_\nu / \varkappa_\nu\,. \tag{16.13}$$

Here the value of σ_ν is determined by the *pressure effects*, among other things. Thus, even if the temperature compensation for the difference in degree of ionisation between giants and dwarfs were complete*, there would still necessarily be a definite difference between the intensities of absorption lines in their spectra.

It follows from all we have said that, when we consider stars of the same spectral class but with different g, and measure the equivalent width W_λ, in the spectra of these stars, of any line which is sufficiently sensitive to a change in p_e (and thus in g), we should expect a definite relation between g and W_λ. This is wholly confirmed by the appropriate observations. However, in practice we establish not the relation between W_λ and g, but between W_λ (or some other characteristic of the line strength, such as an eye estimate of the strength) and the absolute magnitude of the star or its bolometric luminosity L. This relation, as is well known, is the basis of the method of spectroscopic parallaxes. The possibility of passing from the relation between W_λ and g to that between W_λ and L depends on the existence of the mass-luminosity law (see below). This law, although not universal, is nevertheless obeyed by fairly large groups of stars in the Hertzsprung-Russell diagram. The method of spectroscopic parallaxes is generally applied within such groups of stars. As is well known, it consists essentially in the following. Having selected for a given spectral class two lines whose intensity ratio changes noticeably with the absolute magnitude of the star, we determine this ratio for a number of stars of known† and different absolute magnitudes (but of the same class). A graph is next constructed: on the axis of abscissae is placed the absolute magnitude, and on the axis of ordinates the ratio in question. Using this calibrated graph, we can determine the unknown absolute magnitude of any star (again of the same class), if we have measured the ratio of intensities of the lines concerned, in its spectrum. For greater accuracy we can use not one pair of lines, but several.

It is quite evident that the array of pairs of lines necessary to determine the absolute magnitude varies along the spectral sequence. Moreover, the array of pairs of lines and the corresponding calibrated graphs

* That is, if the decrease of p_e in the giant were compensated for all atoms by a decrease in its temperature relative to the dwarf.

† Determined, say, from the trigonometrical parallax.

(see above) are different for instruments with different dispersions. For early-type stars, the absolute magnitude can be determined simply from the equivalent widths of the lines of the Balmer series. We have already explained the corresponding method in Chapter 13. The reader is referred to the literature for a list of the lines used in the method of spectroscopic parallaxes, and for details of the method.

It must be remarked that the purely empirical method of spectroscopic parallaxes has given very good results in practice. It is based, as we have just said, on the assumption that all stars of the spectral class concerned and neighbouring classes follow the "mass-luminosity law":

$$L = \Phi\,(\mathsf{M})\,. \tag{16.14}$$

We then find from formulae (4.22) and (13.28) that the acceleration g due to gravity must be a function of the luminosity and of the temperature T_e:

$$g = F\,(L,\,T_e)\,. \quad (16.15$$

On the other hand, as we have seen above, the spectral class Sp depends on T_e and (to a lesser extent, it is true) on g. Hence we can write

$$Sp = f\,(g,\,T_e)\,. \tag{16.15'}$$

Having found T_e from (16.15′) as a function of Sp and g and substituted in (16.15), we find

$$g = \phi\,(L,\,Sp)\,, \tag{16.15''}$$

from which it follows that, for a given Sp, the acceleration g due to gravity is a function of L, and conversely. This is why the value of W_λ, for a given Sp, must depend mainly on L.

The dependence of W_λ on g for a given T_e (or for a given Sp) is called the **absolute-magnitude effect**. This effect was the object of particular study in the "theory of ionisation". However, all the predictions of this theory regarding the dependence of W_λ on g must now be entirely revised. Besides its schematic nature, this theory took no account of the line broadening caused by collision damping. The efficiency of this process is determined by the density of matter, and must consequently depend on the value of g, and directly on the absolute magnitude when T_e is constant.

In order to investigate the absolute-magnitude effect theoretically, it is necessary to proceed in the manner discussed above, namely to calculate the structure of the photospheres of stars for various T_e and g, and to construct, on the basis of these calculations, the contours of the absorption lines in which we are interested. Then, having decided on a definite value of T_e (or Sp), we can study the dependence of W_λ on g.

It must be remarked that such calculations are restricted as regards stars of high luminosity, since calculations based on "dynamical" values of g are impossible for the atmospheres of supergiants (see Chapter 13). Furthermore, the broadening of lines by the effects of turbulence has a distorting effect in the atmospheres of supergiants (Chapters 12 and 13).

The present state of the theory of the absolute-magnitude effect is in general very unsatisfactory. A great amount of work remains to be done here. As yet the whole problem has been developed in a purely empirical manner.

Besides the purely quantitative method of spectroscopic parallaxes, the so-called **luminosity classes** have recently come into use. The luminosity classes are denoted by Roman figures. Stars of class I are supergiants, while those of class V are chiefly stars of the main sequence. Fig. 45 shows the corresponding diagram, which gives the preliminary

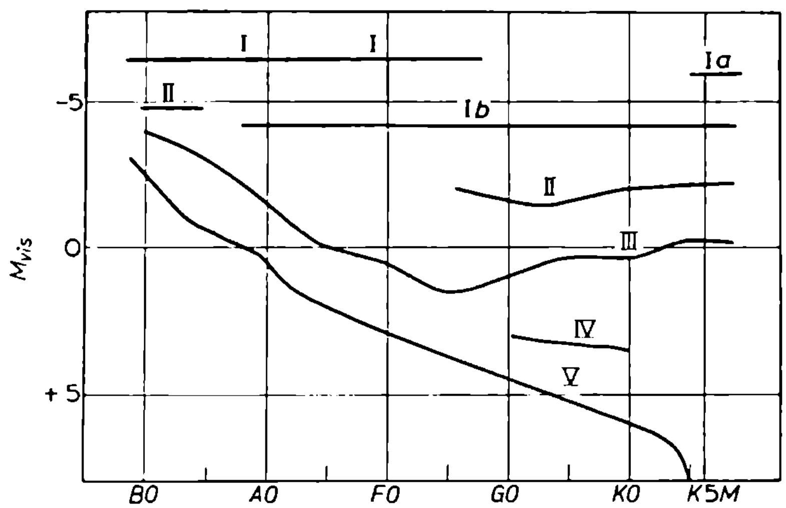

Fig. 45

calibration of the luminosity classes [e. g. 92]. On the axis of abscissae is placed the spectral class, and on the axis of ordinates the absolute stellar magnitude.

In connection with the interpretation of the spectral classification, it must be remarked that it will, of course, be necessary in future to refine the classification continually, in order to take account of various supplementary factors which characterise the spectra of stars: turbulence, deviations from the "normal" chemical composition, etc.

5. **The scale of effective temperatures.** In conclusion, we shall briefly consider the subject of the **scale of stellar effective temperatures** [97]. By this we shall understand the relation, established from observation, between the spectral class of a star and its effective temperature. It is clear that, for stars of different luminosity, this relation will be different, because of the considerations given above.

We shall be interested only in effective temperatures, since we have seen that the effective temperature of a star is in fact the basic temperature parameter. According to the relation (4.22), it determines the total, i. e. bolometric, luminosity L of the star, and by (4.21) the flux πH, which, together with the acceleration g due to gravity and the chemical composition, determines the temperature distribution in the stellar atmosphere and the spectrum of the star. Other kinds of stellar temperature are parameters which usually play only a restricted part.

The most direct method of determining T_e is to use the fundamental formula

$$L = 4\,\pi\,R^2 \sigma\,T_e^4\,, \tag{16.16}$$

which connects L, R and T_e and gives the definition of T_e. However, we here encounter a difficulty (insuperable without special investigations), in that the Earth's atmosphere cuts off the part of stellar ultra-violet radiation with wavelengths $\lambda < 3000$ Å. For stars with fairly high T_e, the proportion of energy not transmitted by the Earth's atmosphere may considerably exceed the proportion of energy which reaches the observer. Only for stars cooler than the Sun and the Sun itself do the spectral regions cut off by the Earth's atmosphere play a very small part. Hence, for stars of different spectral classes, the methods of determining T_e must be modified in some way, and sometimes we must resort to indirect methods.

For very hot stars (hotter than those of class B 2), the most reliable methods of determining T_e are so far those concerned with the study of absorption (or emission) lines. A comparison of the theoretical energy distribution in the continuous spectrum of these stars with the distribution found from observation leads to considerably more uncertain results, since at high temperatures the relative distribution of energy in the observable part of the spectrum depends only slightly on the temperature.

The most reliable method of determining T_e from absorption lines for hot stars is the "maximum" method of the theory of ionisation, which we have considered in the present chapter. The values of T_e obtained by this method are shown in Table 9. The excitation potential refers to the lower level of the line whose maximum is given from observation. The spectra in the two columns are given according to the results of different observers.

Three methods can be used for stars of classes B 2 to A 0:

(1) the method of maxima just discussed;

(2) a comparison of the theoretical and observed energy distribution in the continuous spectrum of the stars;

(3) the use of bolometric corrections.

Let us consider the second of these methods. The comparison of the theoretical curves with those obtained from observation should, in order to achieve the greatest accuracy, include as wide a range of wavelengths as possible, among them the discontinuity at the limit of the Balmer series and the interval from $\lambda = 3000$ Å to this limit ($\lambda = 3646$ Å). Having selected the theoretical curve, constructed for

Table 9 [based on **117**]

Element	Ionisation potential	Excitation potential	T $\log_{10} g = 4{\cdot}4$	T $\log_{10} g = 2{\cdot}4$	Spectrum	
Si II	16·3	9·3	12,900°	11,200°	A 0	
C II	24·3	18·0	21,000	16,800		B 3
N II	29·5	18·4	21,000	18,700	B 2	B 1—2
Si III	33·3	18·9	22,000	19,400	B 1—2	B 1—2
He I	24·5	20·9	18,000	14,400	B 2—3	
O II	34·8	23·3	24,000	21,000	B 1	B 1
Si IV	44·9	23·9	31,500	27,000	O 9	O 9
N III	47·4	27·3	35,000	31,500	O 7—8	O 9
C III	47·7	29·4	30,600	28,000	O 9 ?	O 9—B 0
O III	54·9	33·7	37,400	33,600	O 9—B 0	O 8
N IV	77·3	39·7	39,000	33,600		
He II	54·2	50·8	44,000	39,000		
C IV	64·2	55·5	50,000	44,000		

some temperature T_e', which best matches the observed energy distribution in the spectrum of a star of the given class, we can suppose that this temperature is characteristic of the spectral class concerned. This method is quite reliable, since the energy distribution beyond the limit of the Balmer series, that up to the limit, and the magnitude of the discontinuity D at the series limit [see formula (6.28)] can be considered as independent characteristics of the spectrum from the observational point of view.

Let us now consider the third method. From the constructed theoretical energy distribution in the continuous spectrum, we can calculate the theoretical bolometric corrections for stars of various classes:

$$\Delta M_{\text{bol}} = M_{\text{bol}} - M_{\text{vis}} . \tag{16.17}$$

Then, knowing the absolute visual (or photovisual) magnitude M_{vis} of the star, we can determine L from the obvious formula

$$2{\cdot}5 \log_{10} (L/L_{\odot}) = (M_{\text{bol}})_{\odot} - M_{\text{bol}} . \tag{16.18}$$

[We shall discuss the values of $L_{\odot}$ and $(M_{\text{bol}})_{\odot}$ later.] Thus the problem reduces to a determination of the radius of the star, followed by finding the value of T_e from the known L and R, in accordance with (16.16).

The value of R can be directly determined only in two cases. (1) If the star belongs to an eclipsing variable system and the spectra of the two components can be distinguished (otherwise it would not be possible to determine the spectral class to which the value of T_e refers). Further, the parallax of the system must be known. In this case the value of R is found from calculations of the orbit. (2) If the distance of the star is known, interferometric measurements can be used; these give the angular radius of the star. There are as yet no such measurements for early-type stars, because of their relatively small dimensions.

The third method which we have considered for determining effective temperatures is really a refinement of the temperatures obtained by the second method, since here it is necessary to use an already known theoretical energy distribution in the continuous spectrum of the stars.

The scale of effective temperatures for stars hotter than A 0 stars [97], constructed by G. P. Kuiper by the methods given above, is shown in Table 10.

Table 10

Spectrum	$\log_{10} T_e$	Δm_b	Spectrum	$\log_{10} T_e$	Δm_b
O 8	4·60	—3·8	B 3	4·27	—2·03
O 9	4·50	—3·24	B 4	4·23	—1·82
B 0	4·40	—2·70	B 5	4·19	—1·60
B 1	4·39	—2·48	B 6	4·16	—1·44
B 2	4·31	—2·24	B 8	4·09	—1·07
			A 0	4·03	—0·76

The subject of the temperature scale for stars of the earliest classes is very indefinite. An extrapolation made by Kuiper gave very high effective temperatures for these stars. For the class O 5, the value of T_e was 80,000°, for the class O 6 63,000°, and for the class O 7 50,000°. However, there are a number of more recent investigations which suggest that these temperatures are greatly exaggerated. It appears that stars of the class O 5 have effective temperatures not exceeding 45,000° or 50,000°, while the general change of T_e with spectral class is slower for O-type and early B-type stars than that shown in Table 10. This subject requires very careful investigation.

The scale of bolometric corrections for B 2 to A 0 stars was prepared by É. R. Mustel' from theoretical curves like Fig. 17. For hotter stars, the bolometric corrections have been calculated from a simplified theory of radiative equilibrium (A. Pannekoek's calculations).

For stars of the classes A to F, the application of the methods mentioned for A 0 to B 2 stars is troublesome, because the absorption of radiation in their atmospheres is determined simultaneously by hydrogen

and by H ions. There are as yet no reliable theoretical calculations of the flux πH_λ for stars of these classes. Hence we must here resort to interpolation methods, since, for stars hotter than those of the class A 0 or cooler than those of the classes G 0 to F 8, the relation between the effective temperature and the spectral class becomes fairly reliable. Hence we now pass immediately to stars cooler than those of the classes G 0 to F 8.

Here matters are very considerably simplified by the fact that, in the spectra of F 8 to G 0 stars, the proportion of energy which belongs to the inaccessible region with $\lambda < 3000$ Å is small, and for stars of the classes K and M it is altogether negligible. The values of L for this group of stars can be found by means of radiometric measurements, which give the values of Δm_b. Knowing Δm_b, we can determine L. A scheme of values of Δm_b was first constructed, using this method, by P. P. PARENAGO [118], and later by KUIPER. The radius R of the star can be determined, say, by the methods given above. In particular, interferometric measurements exist for some red giants.

For dwarfs, we must use eclipsing binary stars which show two distinct spectra. The most reliable temperature value is, of course, that of the Sun, a dwarf of class dG 3. The effective temperature of the Sun, determined from the solar constant, is 5710°, $(M_{\text{bol}})_\odot = +4^{\text{m}}{\cdot}7$, and $L_\odot = 2 \times 10^{33}$ ergs/sec. The scheme of effective temperatures of stars of classes G, K, and M, determined in the manner described, is given in Table 11. The temperatures of stars of classes A 5, F 0 and F 5 are found, as stated above, by interpolation, taking into account the information which can be obtained from the intensities of absorption lines.

Table 11

Spectrum	T_e	Spectrum	T_e	Spectrum	T_e	Spectrum	T_e
A 0	10,700°	dG 0	6000°	gG 0	5200°	gM 0	3400°
A 5	8530	dG 2	5710			gM 2	3200
F 0	7500	dG 5	5360	gG 5	4620	gM 4	2930
F 5	6470	dK 0	4910	gK 0	4230	gM 6	2750
		dK 2	4650			gM 8_e	2590
		dK 5_+	3900	gK 5	3580		
		dM 2	3200				

From Table 11 we can estimate the difference between T_e for a dwarf and T_e for a giant of the same class.

The scale of bolometric corrections, obtained by the method given above, for stars of various classes is given in Table 12. In this table, M is the absolute visual magnitude of the star; a colon after the result

indicates that it is uncertain. Values in brackets were obtained by extrapolation.

The tables given here are provisional. The temperatures of very hot stars and of those of classes A and F are the most uncertain. However, the accuracy of the results in these tables is quite sufficient for many practical problems.

Table 12

Spectrum	Main sequence	$M = 0{\cdot}0$	$M = -4{\cdot}0$	Spectrum	Main sequence	$M = 0{\cdot}0$	$M = -4{\cdot}0$
F 2	—0·04	—0·04	—0·04	K 4	—0·55	—1·11	—1·56
F 5	—0·04	—0·08	—0·12	K 5	—0·85	—1·35	—1·86:
F 8	—0·05	—0·17	—0·28	K 6	—1·14		
G 0	—0·06	—0·25	—0·42	M 0	—1·43	—1·55	—2·2
G 2	—0·07	—0·31	—0·52	M 1	—1·70	—1·72	—2·6
G 5	—0·10	—0·39	—0·65:	M 2	—2·03	—1·95	—3·0:
G 8	—0·10	—0·47	—0·80:	M 3	[—2·35]	—2·26	—3·6:
K 0	—0·11	—0·54	—0·93:	M 4	[—2·7]	—2·72	
K 2	—0·15	—0·72	—1·20	M 5	[—3·1 :]	—3·4:	
K 3	—0·31	—0·89	—1·35				

The scale of effective temperatures and bolometric corrections can be used, in particular to draw lines of equal radius on the "spectrum-luminosity" diagram, and if the masses of the stars are known, then lines of equal mass and equal density can be drawn also.

Fig. 46 shows such a diagram, constructed by P. P. PARENAGO. It is very useful, both for practical problems of stellar astronomy and for preliminary calculations in various astrophysical problems. On the axis of ordinates are placed the absolute bolometric magnitudes of the stars, and on the axis of abscissae $\log_{10} T_e$ and the spectral classes. The vertical thick lines in the diagram mark the existing groups and sequences of stars. Their height gives the approximate dispersion of the magnitudes M. The sequences and groups shown in the diagram are as follows: *AB*, the main sequence; *CD*, the giants; *EF*, the supergiants; *GH*, the cepheids; *JK*, the subgiants; *L*, the white dwarfs; *MN*, the subdwarf sequence discovered by P. P. PARENAGO; *OP*, the white-blue sequence whose continuity was discovered by B. A. VORONTSOV-VEL'YAMINOV.

The level lines on the diagram are shown as follows. The continuous lines are lines of equal radius, calculated from the formula

$$\log_{10} R = 8{\cdot}50 - 0{\cdot}2\, M_{\text{bol}} - 2 \log_{10} T_e\,. \qquad (16.19)$$

This formula is obtained from formula (16.16), taking into account that the absolute bolometric magnitude of the Sun $(M_{\text{bol}})_\odot = +4{\cdot}7$. The

dashed lines are lines of equal mass. They were obtained by P. P. PARENAGO from his own investigations and from those he conducted together with A. G. MASEVICH.

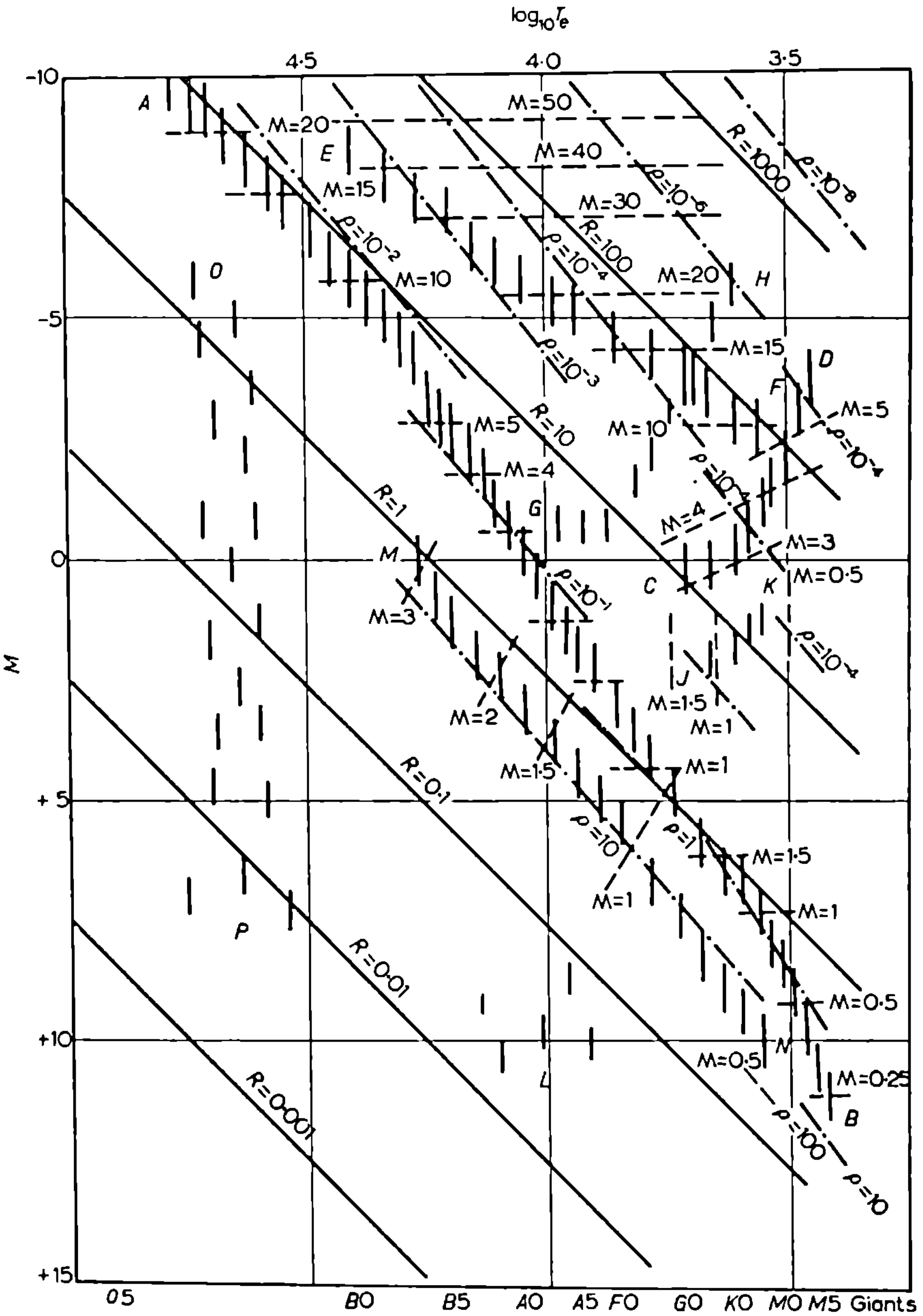

FIG. 46

We now give the various empirical formulae for the relation between L_{bol} and the mass M of the star. In the upper part of the main sequence, for stars hotter than the Sun,

$$M = 0{\cdot}89\, L_{\text{bol}}^{0{\cdot}25} . \tag{16.20}$$

For the remaining stars in the main sequence,

$$\mathsf{M} = 2{\cdot}4\, L_{\mathrm{bol}}{}^{0{\cdot}44} \,. \tag{16.21}$$

For the subdwarfs,

$$\mathsf{M} = L_{\mathrm{bol}}{}^{\frac{1}{4}}\, R^{-\frac{1}{3}} \,. \tag{16.22}$$

For the subgiants,

$$\mathsf{M} = 1{\cdot}8\, L_{\mathrm{bol}}{}^{\frac{1}{2}}/R \,. \tag{16.23}$$

For the giants,

$$\mathsf{M} = 1{\cdot}1\, L_{\mathrm{bol}}{}^{0\,4}\, R^{-0{\cdot}25} \,. \tag{16.24}$$

For the supergiants,

$$\mathsf{M} = 1{\cdot}5\, L_{\mathrm{bol}}{}^{0{\cdot}27} \,. \tag{16.25}$$

The unit of measurement of L_{bol} is $(L_{\mathrm{bol}})_{\odot}$, and of M, $\mathsf{M}_{\odot}$.

The dot-and-dash lines in the diagram show the lines of equal density, which are obtained from the known masses and radii, according to the formula

$$\log_{10} \varrho = \log_{10} \mathsf{M} + 0{\cdot}6\, M_{\mathrm{bol}} + 6 \log_{10} T_e - 25{\cdot}50; \tag{16.26}$$

this is derived from (16.19) and the formula $\mathsf{M} = \frac{4}{3}\pi R^3 \varrho$. We are speaking, of course, of mean densities.

In the present chapter we have considered a two-parameter spectral classification. The situation is really much more complicated. Besides the effective temperature and the acceleration due to gravity at the surface of the star, the nature of the star's spectrum may be determined by other factors also, such as, for example, the rotation of the star, the turbulence of the gases which form its atmosphere, and so on. Moreover, the deviations of the chemical composition of the stellar atmosphere from the "normal" composition may play a considerable part. Finally, the spectra of the various sequences and groups of stars, such as the subdwarf sequence, the group of "metallic-line" stars, etc., demand especial study.

PART III.
THE PHYSICS OF THE SOLAR ENVELOPES

Chapter 17. The structure of the solar photosphere. Granulation. Convection

1. **The structure of the solar photosphere.** It would be possible to determine the structure of the solar photosphere theoretically, by the same methods as we used in Chapter 7 to study the structure of the photospheres of stars of the class A 0. However, as we have said in Chapter 6, the theory of radiative equilibrium, as developed for the solar photosphere, requires correction: the mean absorption coefficient which was introduced to make the theory agree with the observations is 1·4 times the theoretical value. Hence it is best not to apply the theory of radiative equilibrium to the Sun, but to find the temperature distribution in its photosphere directly from observation. This can be done by using the method explained at the end of Chapter 7. We must study, as accurately as possible (in particular, close to the limb) the energy distribution over the solar disc in some wavelength λ, and then this gives the law of variation of the function B_λ with the optical depth τ_λ for the same wavelength, on applying formula (7.31).

Let us now consider how, knowing the relation between B_λ and τ_λ for some wavelength λ, and therefore the relation between T and τ_λ, we can find the depth distribution of the other physical quantities: the total gas pressure p, the electron pressure p_e, the density ϱ, etc. We write the equation of hydrostatic equilibrium; the results of Chapter 7 show that we can neglect the part played by radiation pressure:

$$\mathrm{d}p = g\,\varrho\,\mathrm{d}h\,. \tag{17.1}$$

For the parts of the spectrum between $\lambda = 4000$ Å and $\lambda = 10{,}000$ to 15,000 Å, in which the relation between B_λ and τ_λ is usually found, the absorption of radiation is determined mainly by negative hydrogen ions. In this case the absorption coefficient, referred to 1 gram of photospheric matter, is given by formula (13.13), where the value of $k_\lambda(\mathrm{H}^-)$ can be taken from Fig. 15. Furthermore, as we said in Chapter 13, we can take $(n_0)_\mathrm{H} \approx n_\mathrm{H}$ throughout the solar photosphere. Hence we can write

$$\mathrm{d}\tau_\lambda = \varkappa_\lambda(\mathrm{H}^-)\,\varrho\,\mathrm{d}h = n_\mathrm{H}\,k_\lambda(\mathrm{H}^-)\,p_e\,\mathrm{d}h\,. \tag{17.2}$$

Dividing (17.1) by $n_H\, k_\lambda(H^-)\, p_e\, dh$ and using (13.9), we obtain

$$\frac{dp}{d\tau_\lambda} = \frac{g\, m_H\, b}{k_\lambda(H^-)\, p_e}, \tag{17.3}$$

where b is the quantity in braces in (13.9) and is equal to about 1·8. If, however, the content a_s of helium in the Sun's atmosphere is the same as in the atmospheres of the early-type stars (see Chapter 15), where apparently $a_s \simeq 0{\cdot}05$, then b will be 1·2. We notice here that $k_\lambda(H^-)$ is a function of temperature and wavelength (or frequency) only. Multiplying (17.3) by $2\,p$, we have

$$\frac{dp^2}{d\tau_\lambda} = \frac{2\, g\, b\, m_H}{k_\lambda(H^-)\, p_e/p}. \tag{17.4}$$

We shall now obtain an expression for the ratio

$$p_e/p = n_e\, kT/n\, kT = n_e/n\,, \tag{17.5}$$

where n_e is the number of free electrons in 1 cm^3, and n the total number of particles in 1 cm^3. For n_e we can write:

$$n_e = n_H\, x_H + \sum n_s\, x_s = n_H(x_H + \sum a_s\, x_s)\,, \tag{17.6}$$

where a_s is the content of the sth element relative to hydrogen [formula (5.1)]. The summation must, strictly speaking, be extended over all elements. However, in practice we can restrict ourselves to the metals, and then only to those for which the value of $a_s\, x_s$ is comparatively large. These are chiefly the metals with large relative contents a_s and relatively low ionisation potentials (Mg, Na, Ca, Al, Si, Fe). The ionisation of other elements (He, O, C, N, etc.) gives very few free electrons in the solar atmosphere. The second ionisation can also be neglected here.

The total number of all particles in 1 cm^3 of the solar photosphere is made up of atoms and electrons. Only hydrogen and helium among the atoms need be taken into consideration, since according to Table 1 there are considerably fewer atoms of other elements. The number of electrons in 1 cm^3 is determined by formula (17.6). Here it is clear that $n_e \ll n_H + n_{He}$, since $x_H \ll 1$ in the solar photosphere, and the number of electrons in 1 cm^3 produced by the ionisation of metals cannot be greater than the number of metal atoms themselves, which in turn are considerably fewer than the hydrogen atoms. Hence we can write approximately

$$n \approx n_H + n_{He} = n_H\,[1 + a_s(He)]\,. \tag{17.7}$$

The expression (17.5) can now be rewritten

$$\frac{p_e}{p} = \frac{n_e}{n} = \frac{x_H + \sum a_s\, x_s}{1 + a_s(He)} = \Phi(p_e, T)\,, \tag{17.8}$$

since the quantities x_{H} and x_s are functions of p_e and T only and can be calculated from the ionisation formula (5.16)*. It is clear that to calculate $\Phi(p_e, T)$ the values of a_s must be given, for example from Table 1.

The equation (17.4) with the relation (17.8) is solved by the iterational method. The integration starts at the boundary of the photosphere, where $\tau_\lambda = 0$. For this value of τ_λ we have

$$T = T_0, \ p_e = (p_e)_0 \,. \tag{17.9}$$

Without loss of accuracy in the values of p and p_e obtained by integrating (17.4), we can assume that $(p_e)_0 = 0$, say, though the value of p_e at the boundary of the photosphere *proper* is of course different from zero, and somewhat exceeds the value of p_e at the base of the chromosphere (see Chapter 21). As a result of integrating (17.4), using (17.8), we obtain the relation between the quantities T, p and p_e, while the relation between T and τ_λ is known from observation.

Let us now consider how we can find the distribution of the other quantities in the solar photosphere. We turn first to the total density of matter ϱ. In accordance with (17.7) we have $p = n\,kT \approx [1 + a_s(\mathrm{He})]\, n_{\mathrm{H}}\, kT$, while n_{H} is related to the density ϱ by (13.9). Consequently, knowing p and T, we can determine ϱ. Finally, the transition to linear depths is made by means of (17.1):

$$h - h_0 = \frac{1}{g} \int \frac{\mathrm{d}p}{\varrho} \,, \tag{17.10}$$

where h_0 is a constant of integration. The integration in (17.10) is carried out numerically, the relation between p and ϱ having already been established.

Table 13 was prepared by V. S. Berdichevskaya from the results of P. ten Bruggencate, H. Gollnow, and F. W. Jäger [165]. These results (the relation between B_λ and τ_λ) were obtained from a study of the darkening to the limb of the solar disc, made at the time of a partial eclipse. The wavelength $\lambda = 5010$ Å.

A detailed and critical discussion of the various models of the solar photosphere that have been proposed by different authors has recently been given by M. Minnaert [77, Chapter 3]. From these models, he has constructed a composite model of the photosphere (Table II of his article), which takes account of the most reliable properties that have been established for the various models. It is possible that this model is closer to reality than the one given in Table 13, although

* Some authors introduce a mean degree of ionisation of the metals x_{M}, which is equal to the sum $\sum a_s\, x_s$.

Table 13

h km*	τ_λ	$\frac{5040}{T}$	T	$a(H^-)$	$\log_{10}p_e$	$\log_{10}p$	$\log_{10}n_H$	$\bar{\tau}$	$T(h)$	$\log_{10}p_e$	$\log_{10}p$
	for $\lambda = 5010$ Å							for the mean absorption coefficient			
	0·00	1·143	4410	$14{\cdot}1 \times 10^{-26}$							
−223	0·005	1·090	4623	11·7	$\bar{1}$·71	4·11	16·24	0·00	4670	$\bar{1}$·84	4·18
−188	0·01	1·053	4786	10·0	1·91	4·27	16·38	0·02	4690	1·98	4·35
−152	0·02	1·029	4898	9·1	0·10	4·44	16·53	0·04	4740	0·16	4·52
−132	0·03	1·011	4985	8·5	0·22	4·53	16·61	0·05	4780	0·26	4·61
−104	0·05	0·984	5122	7·6	0·38	4·65	16·72	0·09	4860	0·38	4·75
− 94	0·06	0·963	5234	7·0	0·48	4·69	16·75	0·10	4900	0·44	4·79
− 79	0·08	0·946	5328	6·5	0·59	4·75	16·81	0·14	4970	0·50	4·86
− 69	0·10	0·935	5390	6·2	0·66	4·79	16·84	0·16	5025	0·58	4·90
− 67	0·11	0·929	5425	6·1	0·70	4·80	16·85	0·18	5045	0·60	4·91
− 61	0·12	0·927	5437	6·0	0·72	4·82	16·87	0·20	5085	0·64	4·94
− 57	0·13	0·927	5437	6·0	0·73	4·84	16·89	0·22	5115	0·67	4·96
− 52	0·14	0·929	5425	6·1	0·74	4·86	16·91	0·23	5160	0·70	4·98
− 49	0·15	0·937	5379	6·2	0·71	4·87	16·92	0·25	5180	0·72	4·99
− 22	0·24	0·920	5478	5·8	0·85	4·98	17·02	0·40	5410	0·94	5·11
− 8	0·31	0·896	5625	5·3	1·01	5·04	17·06	0·55	5600	1·10	5·16
+ 4	0·40	0·874	5765	4·8	1·16	5·08	17·09	0·70	5770	1·28	5·21
+ 14	0·48	0·855	5895	4·4	1·28	5·12	17·12	0·90	5970	1·48	5·25
+ 26	0·65	0·821	6138	3·7	1·53	5·16	17·15	1·34	6350	1·84	5·29
+ 33	0·81	0·796	6332	3·3	1·71	5·18	17·16	(1·55)	(6550)	(2·0)	(5·32)†
+ 39	0·98	0·775	6503	3·0	1·87	5·20	17·17				
+ 44	1·13	0·758	6649	2·8	2·00	5·22	17·18				
+ 47	1·29	0·742	6792	2·6	2·12	5·23	17·18				
+ 50	1·46	0·728	6923	2·4	2·23	5·24	17·18				
+ 52	1·60	0·716	7039	2·2	2·32	5·25	17·18				
+ 56	2·00	0·690	7304	2·0	2·52	5·26	17·18				
+ 61	2·5	0·683	7380	1·9	2·58	5·27	17·19				

* $h = 0$ for the layer where $T = 5713°$; $h > 0$ towards the centre of the Sun, and $h < 0$ towards the boundary of the photosphere.

† Values in parentheses are extrapolated.

there still remain a number of points to be settled. In particular, we do not know the exact content of helium (see above).

It should also be noted that the data at present existing on the law of variation of the intensity of radiation over the solar disc are affected by various errors and differ somewhat among themselves. (Further very careful observations are necessary, particularly at the limb of the disc.) Nevertheless, even with only these still imperfect data, the method in question is very much superior to that which uses the theory of radiative equilibrium.

We could have solved the same problem of the Sun's photosphere by the method employed in Chapter 7 for stars of the class A 0. Here it is necessary to use the theoretical law of temperature distribution in the solar photosphere. According to what we said about the solar photosphere in Chapter 6, we can take this law from (4.26), introducing, of course, the mean continuous absorption coefficient [formula (6.30)]. The results of such calculations* are given in the last four columns of Table 13. Naturally, without these results we cannot go on to study the greater optical depths, from which radiation does not reach us, and so the relation between B_λ and τ_λ cannot be derived.

2. **Convective processes in the solar photosphere. Granulation.** The model of the solar photosphere which we have considered is, of course, a somewhat schematic one. It is essentially a model of a purely static photosphere. However, the presence of *granulation* in the solar photosphere shows that there is a movement of gas, even though the solar photosphere is in equilibrium as a whole. We recall that **granulation** signifies the "graininess" of the observed outer surface of the Sun. The separate luminous "grains" on the darker background are called **granules**. The average dimensions of a granule, according to the investigations of A. P. GANSKIĬ (at Pulkovo) are of the order of 1″ to 2″, which is equivalent to 700 to 1400 km on the Sun's surface. It has been established by observation that the granular structure of the solar surface — the *granulation pattern* — undergoes relatively rapid changes. The mean lifetime of a photospheric granule is of the order of 3 minutes.

The general form of the granulations, and the manner of the appearance and disappearance of granules, leads us to suppose that we here observe *rising and falling masses of gas.* Recently conducted spectroscopic investigations of granulation confirm this view [125]. A value of about 0·4 km/sec is obtained for the mean velocity of radial motion v_0 of the granule, taken in the sense of formula (11.26). This is considerably less than the turbulent velocities obtained from the curve of growth

* The data in these columns were calculated by V. S. BERDICHEVSKAYA on the basis of G. MÜNCH's model of the solar photosphere [93].

for the Sun, which amount to about 1·5 km/sec. It is quite possible that this difference is connected with the difference in the dimensions of the turbulence elements. However, the situation is not yet clear, since the nature of the "turbulent" velocities found from the curve of growth may prove to be different from that of the velocities of the granulation currents. Moreover, further very careful examination of the observations of the velocities themselves is necessary (in measuring the velocities of granulation currents, the effect of scattered light must be taken into account, and so on). In particular, G. THIESSEN finds that, when scattered light is taken into account, the above value of 0·4 km/sec is increased to 1·8 km/sec.

However, despite these uncertainties, we must suppose that rising and falling currents are present in the photosphere, i. e. that the latter is *internally unstable*. Moreover, it must be borne in mind that the properties of the granulation (granule dimensions, brightness contrast, etc.) are *independent of the heliographic latitude* (on the Sun's surface) *and of the 11-year period of variation of the solar activity.*

We have already said that the model of the solar photosphere considered above (Table 13) is one of a purely static photosphere. However, the calculations that have been performed, especially if they are extended to large values of τ_λ, show that the lower layers of the solar photosphere cannot possibly be in stable equilibrium. There must exist *internal convective currents* in them, which can with a fair amount of justification be identified with the phenomena of granulation. (These currents must be turbulent in nature, as is shown by calculation.)

Starting from the idea that the solar photosphere (or at least its outer layers) is in radiative equilibrium, let us consider in which cases the convective currents in such an atmosphere can be stable, and in which cases they cannot. Let an element of gas in the photosphere acquire, for some reason, a momentum directed outwards. It will then begin to rise and expand. Its temperature will thereby change in a continuous manner. We shall now make the following assumptions: (1) the expansion process takes place adiabatically, (2) the viscosity forces which might retard the element in question can be neglected. Then, in order that the motion of the element should not be stopped, it is necessary that the *adiabatic temperature gradient* (the change with depth of the temperature of an adiabatically expanding and rising element) should be less than the gradient corresponding to radiative equilibrium. This is easily understood from Fig. 47. The upward motion of the element corresponds to a movement on the graph along the line Oa. Along the whole path of the element, its temperature is higher than that of the solar photosphere surrounding it, and consequently the motion

of the element will be continually maintained. Thus the convective currents which arise in the photosphere will persist if the inequality

$$\left(\frac{dT}{dh}\right)_{ad} < \left|\frac{dT}{dh}\right|_{rad} \tag{17.11}$$

is satisfied. If the contrary inequality holds (see dashed curve in Fig. 47), then the motion of the element will be immediately stopped, since it is cooler and denser than the surrounding medium.

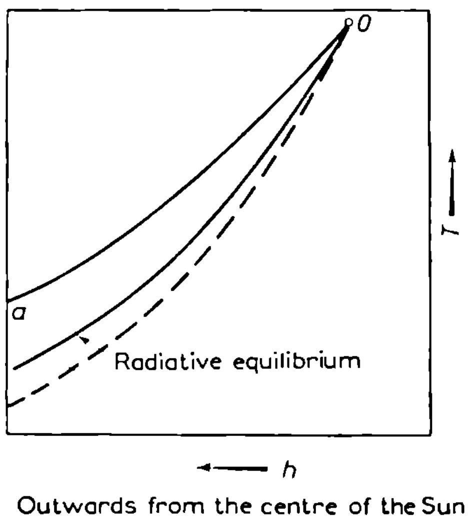

FIG. 47

Thus, if the inequality (17.11) is satisfied, any small variations of local temperature lead to the appearance of ascending and descending currents which, as follows from Fig. 47, will tend to alter the temperature gradient of the medium.

In order to study the convection which occurs when the inequality (17.11) holds, we shall consider currents so weak that they do not destroy the general equilibrium of the atmosphere. In other words, we shall suppose that as the element moves upwards or downwards *the pressure exerted by the element on the surrounding photosphere* is equal to *the pressure exerted on the element by the surrounding photosphere*. The state of the medium when these conditions are fulfilled is called convective equilibrium [28, p. 84].

Since equilibrium is maintained, we can apply the formula of hydrostatic equilibrium (17.1), which, using the equation of state of the gas

$$p = \frac{R}{\mu} \varrho T, \tag{17.12}$$

takes the form

$$\frac{dp}{dh} = \frac{g\mu}{RT} p. \tag{17.13}$$

Multiplying both sides of (17.13) by dT/dp, we obtain

$$\frac{dT}{dh} = \frac{g\mu}{R} \frac{d\log_e T}{d\log_e p}. \tag{17.14}$$

Applying this for radiative and for convective equilibrium, we find, instead of (17.11),

$$\left(\frac{d\log_e T}{d\log_e p}\right)_{ad} < \left(\frac{d\log_e T}{d\log_e p}\right)_{rad}. \tag{17.15}$$

In order to find whether the inequality (17.15) is satisfied in the photosphere of the Sun, we must calculate separately the two sides of (17.15).

We take first the left-hand side of the inequality. For adiabatic processes, as is well known, $p^{1-\gamma} T^{\gamma}$ = constant, and consequently

$$\left(\frac{\mathrm{d} \log_e T}{\mathrm{d} \log_e p}\right)_{\mathrm{ad}} = \frac{\gamma - 1}{\gamma}, \tag{17.16}$$

where $\gamma = c_p/c_v$ is the ratio of specific heats of the gas at constant pressure and at constant volume. For a monatomic gas, $\gamma = 5/3$, and so

$$\left(\frac{\mathrm{d} \log_e T}{\mathrm{d} \log_e p}\right)_{\mathrm{ad}} = 0{\cdot}40 .$$

A comparison of this number with the values of $(\mathrm{d} \log_e T/\mathrm{d} \log_e p)_{\mathrm{rad}}$ obtained from Table 13 makes it possible to draw the following conclusion. Down to the level at optical depth $\tau_\lambda \approx 0{\cdot}5$ (for $\lambda = 5010$ Å) the inequality $(\mathrm{d} \log_e T/\mathrm{d} \log_e p)_{\mathrm{rad}} < 0{\cdot}4$ holds, but for $\tau_\lambda > 0{\cdot}5$ the converse, i. e. the inequality (17.15), is true. Thus, *below the level in the solar photosphere where $\tau_\lambda \approx 0{\cdot}5$, there must lie an unstable convective region.*

In order to form an idea of the lower limit of the convective zone, it is necessary to take into account another important circumstance, namely that the ionisation of the atoms varies as the gaseous mass rises or falls. The electrons detached by ionisation must receive a kinetic energy of $\frac{3}{2} kT$. On the other hand, a gas whose degree of ionisation varies behaves approximately like a *polyatomic gas*. For this reason, the effective γ for such a gas, defined as before by the relation (17.16), must be less than 5/3.

The most important gas in this connection is *hydrogen*. A. Unsöld has shown that, in the zone of a stellar atmosphere where hydrogen is partly ionised, the effective γ is so small that currents must inevitably occur. When this circumstance is taken into account, the part played by the convective zone in the solar photosphere becomes much greater, and so does its thickness.

To perform the corresponding calculations, it is necessary to find the value of $(\mathrm{d} \log_e T/\mathrm{d} \log_e p)_{\mathrm{ad}}$, taking the ionisation of hydrogen into account [168, p. 227]. However, in order to use this expression, which of course depends on the pressure and the temperature, we must make some assumptions regarding the structure of the convective zone itself. There is reason to suppose that the occurrence of convective motions does not appreciably destroy the radiative equilibrium, so that the latter exists in the convective zone also. In fact, even when convective currents are present, the transfer of energy in the photosphere is effected mainly by radiation, owing to the great rapidity of

radiative exchange. In this case we can use, to find $(d \log_e T/d \log_e p)_{ad}$, the values of p and T obtained for radiative equilibrium, as for instance in Table 13. The linear thickness of the convective zone is then found to be of the order of 700 km.

Unfortunately, the problem of the physical state of the matter in the convective zone is extremely complicated. Thus, the question of the applicability of the formulae of thermodynamic equilibrium (particularly the ionisation formula) to the matter in the convective zone is very obscure. The reason for this is as follows. If some mass of gas moving in the convective zone goes downwards, the degree of ionisation of this mass must always have a tendency to vary, in accordance with the temperature at the level in the atmosphere where the mass is at any given moment. However, it is quite evident that the ability of the mass of gas to take up the degree of ionisation corresponding to its "momentary" level in the atmosphere depends on its rate of motion. If, for example, this rate of motion is very great, then the mass will not be able to change its degree of ionisation sufficiently rapidly, even on moving down to a great depth where the temperature is fairly high. It is true that the actual velocities of the currents in the convective zone are apparently small, about 1 km/sec. However, if photo-ionisation processes only are present (and not collisions), the equilibrium in ionisation (for hydrogen) between the medium and the moving mass of gas is not established even at this velocity. This result is somewhat modified if ionisation by collisions is taken into account. There must nevertheless, apparently, be a definite difference in the ionisation states of currents moving upwards and downwards; the upward-moving currents carry ionised hydrogen to the upper parts of the convective zone, where its recombination may be the cause of the appearance of an excess of high-frequency radiation in the outer layers. If this is so, it means that the unstable zone will be extended, and its upper limit will be raised.

It follows from the above considerations that the cause of granulation phenomena may be seen in the existence of an unstable photospheric zone with $\tau_\lambda > 0{\cdot}5$. The external appearance of the granulation network is also in agreement with these ideas. In fact, if a flat layer of liquid or gas, in which heat transfer takes place by convection, is heated from below, and the layer is well protected from disturbances, a regular network of hexagonal vortices, resembling granulation, is formed in it, the material rising along the axis of each cell and flowing down along its outer surface (the so-called Bénard cells). The distance between the axes of neighbouring convection cells is always about 3·3 times the thickness of the layer. Consequently, if we assume that the thickness of the convective layer is of the order of 700 km, the distance

between the granules will be of the order of 2300 km, and this is not far from reality. Furthermore, since the gaseous masses rising in the convective zone are hotter than the surrounding material, the experiments mentioned above also explain the fact that the granulation consists of bright regions (the centres of the vortex cells), separated by dark interstices. Finally, because the thickness of the convective zone is small in comparison with the radius of the Sun, it is clear why the granulation is independent of the heliographic latitude and of the phase in the 11-year cycle.

It must be pointed out that the phenomenon of granulation is very complex, and needs further careful investigation. In particular, in a recent work by V. A. KRAT [73], based on plates taken under exceptionally good seeing conditions, it is found that the smallest granules have dimensions of the order of $0''\cdot4$, and a contrast of the order of 13 %. Besides the smallest granules, we more frequently find granules of diameter $0''\cdot7$. V. A. KRAT thinks that these granules may be the result of the coalescence of several granules of smaller size. The origin of the larger granules, of diameter $1''$, seems to be the same. By estimating the possible temperature differences between the granules and the intergranular spaces, V. A. KRAT finds that the observed "excitation transfer" effect for atoms in the Sun's "reversing layer" (for lines corresponding to high excitation potentials) can be explained by granulation phenomena.

The general ideas discussed above show that the convective zone must exist not only in the photosphere of the Sun, but also in those of stars. It must certainly arise in all cases where the transition from the neutral hydrogen zone to the ionised hydrogen zone takes place in the photospheric layers.

The depth at which the transitional zone lies must increase as we pass from hot stars to cool ones. In A-type stars of normal luminosity, for instance, the convective zone appears at the outermost layers of the photosphere [20]. In some cases we have even a density reversal, when the density decreases inwards over some region. It is plain that the instability is even more clearly marked in such cases.

Chapter 18. The electrodynamics of the Sun's atmosphere

1. Introductory remarks. The rapid growth of the theory of radiative equilibrium, and the considerable success which it has achieved, have deservedly given it a leading place in theoretical astrophysics. Many observational data can be explained more or less accurately by using

the assumption that the atmosphere is in a state of local thermodynamic equilibrium; where deviations from thermodynamic equilibrium are found, the hypothesis of an anomalously high intensity of solar radiation in the ultra-violet region of the spectrum is employed. The theory of absorption-line formation (see Part II) has more recently achieved notable success also.

However, as methods of observation were perfected (the introduction into practical use of spectroheliographs, spectrohelioscopes, observations of the corona outside eclipse, and especially the slow-motion cinematography of prominences with narrow-band filters which select the H_α line), it became more and more evident that the motions of prominences, and a number of phenomena on the Sun which characterise the solar activity, cannot be reduced to a consideration of radiation alone. Nor were the expected results obtained by taking account of other known forces, of chemical or hydrodynamical origin. Only electromagnetic forces can, even in part, explain the nature of trajectories which recall the lines of force in an electric or magnetic field. Hence the electrodynamics of the solar atmosphere has attracted much attention in recent years and is rapidly being developed.

At present quite a number of theories have been constructed which explain many of the manifestations of solar activity by means of electromagnetic forces. However, the electrodynamics of the solar atmosphere has not yet been completely worked out, and the majority of its conclusions are provisional.

We shall not pause to describe these theories here, but merely consider the physical foundations of the electrodynamics of the solar atmosphere, and some direct applications of it. It is well known that a force

$$\mathbf{f} = e\left(\mathbf{E} + \frac{1}{c}\,\mathbf{v}\wedge\mathbf{H}\right) \tag{18.1}$$

acts on an ion of charge e, moving with velocity $\mathbf{v}$ in an electric field of intensity $\mathbf{E}$ and a magnetic field of intensity $\mathbf{H}$. Here we limit ourselves to the non-relativistic case, i.e. we suppose that $v \ll c$ and $E \ll H$ (in absolute units).

If $\mathbf{E} = 0$, then $\mathbf{f}$ is perpendicular to $\mathbf{v}$, and the absolute magnitude of the velocity does not change. In a constant uniform magnetic field, the motion of the ion is composed of a uniform revolution in a circle with velocity $\mathbf{v}_\perp$ in the plane perpendicular to $\mathbf{H}$, and a *uniform* motion along $\mathbf{H}$ with velocity $\mathbf{v}_{\mathrm{II}}$. It is easily found that the radius of the circle $R = cm_i v_\perp / eH$, and the angular frequency of revolution $\omega_i = eH/m_i c$, where m_i is the mass of the ion.

If $\mathbf{E}$ is parallel to $\mathbf{H}$, $\mathbf{v}_\perp$ does not change, but $\mathbf{v}_{\mathrm{II}}$ increases as it would if the magnetic field were absent.

Since v_{Π} is independent of the presence of the magnetic field and can easily be calculated separately, we shall henceforth consider only $v_{\perp}$, i. e. we shall assume that $\mathbf{v}$ and $\mathbf{E}$ are perpendicular to $\mathbf{H}$. The equation of motion of the ion in this case is

$$m_i \, \mathrm{d}\mathbf{v}/\mathrm{d}t = e\,\mathbf{E} + \frac{e}{c}\,\mathbf{v} \wedge \mathbf{H} \,. \tag{18.2}$$

We introduce a system of co-ordinates moving relative to the old system with velocity

$$\mathbf{U} = \frac{c}{H^2}\,\mathbf{E} \wedge \mathbf{H} \,. \tag{18.3}$$

Then $\mathbf{v} = \mathbf{U} + \mathbf{W}$, where $\mathbf{W}$ is the velocity of the ion in the new system of co-ordinates. Substituting in (18.2) and using the equation $(e/c)\,\mathbf{U} \wedge \mathbf{H} = -\,e\mathbf{E}$, we find that no electric field acts on the ion in the new system of co-ordinates ($\mathbf{E}$ disappears), so that the ion moves in a circle with velocity $\mathbf{W}$. This means that in the old system of co-ordinates the ion moves *in a cycloid* with a *mean velocity* $\mathbf{U}$ which is perpendicular to $\mathbf{E}$, and so the ion does no net work in its motion.

This kind of motion is called **drift**. We notice that the values of the charge and mass of the particle do not appear in the expression (18.3) for the drift velocity. Hence, electrons and various ions move with the same mean velocity ($\mathbf{U}$ is the same for electrons and ions, while $\mathbf{W}$ is different). If, instead of one ion, we place in the same fields a number of ions and electrons forming a *macroscopically neutral plasma*, which is so rarefied that an ion completes a large number of revolutions in the course of its free path, the whole plasma will move with velocity $\mathbf{U}$ and no current will appear in it. If some other force, for example the force of gravity or a pressure gradient, acts on the plasma in place of the electric field intensity, the drift velocities of the electrons and ions are unequal and are in different directions, as is easily shown by replacing $e\mathbf{E}$ in (18.2) and (18.3) by the force $\mathbf{f}$.

Hitherto we have assumed that the magnetic and electric fields are uniform. In non-uniform fields the motions are more complicated; they have been studied in detail by S. A. Boguslavskiĭ. However, in the atmosphere of the Sun, the field may be considered to be practically uniform. The radius of the circle described by a proton moving with a mean thermal velocity $v = 10^6$ cm/sec in a field $H \geq 10^{-2}$ gauss is

$$R \leq \frac{3 \times 10^{10} \,.\, 1{\cdot}7 \times 10^{-24} \,.\, 10^6}{4{\cdot}8 \times 10^{-10} \,.\, 10^{-2}} = 10^4 \text{ cm} \,.$$

The radius of the circle described by an electron is still smaller.

The magnetic field on the Sun cannot change at all noticeably over distances of this order of magnitude, and hence the field within this circle can be supposed uniform. The non-uniformity of the field need be taken into account only for the motions of particles with high energies in a very weak field, as for instance in the theory of cosmic rays.

2. **The conductivity of the solar atmosphere in a magnetic field.** Let us now consider the motion of charges in an actual plasma, where the particles interact with one another. Here we shall suppose that in the intervals between collisions an ion moves freely in accordance with (18.2), while at the moment of a collision it changes its velocity, all directions of the velocity after collision being equally probable, so that the velocity at that moment is zero on the average. For simplicity we shall at first assume that the mean time of free flight τ_1, for the given kind of ion, is independent of the velocity. We shall also suppose that the motions of the ions and electrons take place independently of one another. All these conditions are fulfilled, for example, in a very weakly ionised gas, where the collisions of the electrons and ions take place mainly with neutral atoms, which have an effective elastic collision cross-section $\sigma_0 \approx 10^{-15}$ cm^2. The collisions between ions and electrons are infrequent and do not determine their motions. Let us consider the motion of one kind of ion, of mass m_1 and charge e_1. In the absence of a magnetic field, the acceleration of the ion in the field $\mathbf{E}$ is $d\mathbf{v}/dt = e_1 \mathbf{E}/m_1 = \mathbf{F}_1$, and the mean velocity in the direction of the field is

$$\bar{\mathbf{v}}_1 \approx (e_1 \mathbf{E}/m_1)\, \tau_1 = \mathbf{F}_1\, \tau_1\,. \tag{18.4}$$

This gives a current of density

$$\mathbf{j}_1 = n_1\, e_1\, \bar{\mathbf{v}}_1 = (n_1\, e_1^2\, \tau_1/m_1)\, \mathbf{E}\,. \tag{18.5}$$

The current density is proportional to the field intensity; the proportionality coefficient λ is called the **conductivity**. Since $m_e \ll m_i$, the ionic conductivity is considerably less than the electron conductivity, and that of a plasma in the absence of a magnetic field (λ_0) is practically equal to the electron conductivity:

$$\lambda_0 = \lambda_e = n_e\, e^2\, \tau_e/m_e\,. \tag{18.6}$$

τ_e is determined by the effective cross-section σ_0, the concentration of neutral atoms n_0 and the mean thermal velocity $\bar{v}_e$ of the electrons, thus:

$$\tau_e = \tfrac{3}{4} \cdot \frac{1}{n_0\, \sigma_0\, \bar{v}_e}\,. \tag{18.7}$$

τ_i is similarly determined.

Now let $\mathbf{H} \neq 0$. We take a rectangular system of co-ordinates, with the z axis directed along $\mathbf{H}$, and the y axis parallel to $\mathbf{E}$. The solution of equation (18.2) gives for u and v (the components of the ion velocity $\mathbf{v}_1$ along the x and y axes at time t) the following functions of time and of the velocity at time $t = 0$:

$$\left.\begin{aligned} u &= \left(u' - \frac{F_1}{\omega_1}\right) \cos \omega_1 t + v' \sin \omega_1 t + \frac{F_1}{\omega_1}\,, \\ v &= v' \cos \omega_1 t - \left(u' - \frac{F_1}{\omega_1}\right) \sin \omega_1 t\,. \end{aligned}\right\} \tag{18.8}$$

Here u' and v' are the components of the vector $\mathbf{v}_1'$, the velocity at time $t = 0$. The last term on the right-hand side of the first equation gives the drift velocity. The integration of equations (18.8) gives the co-ordinates of the ion as functions of time and of the initial co-ordinates x' and y':

$$\left.\begin{aligned} x &= x' + \frac{1}{\omega_1}\left\{\left(u' - \frac{F_1}{\omega_1}\right) \sin \omega_1 t + v'(1 - \cos \omega_1 t) + F_1 t\right\}, \\ y &= y' + \frac{1}{\omega_1}\left\{v' \sin \omega_1 t - \left(u' - \frac{F_1}{\omega_1}\right)(1 - \cos \omega_1 t)\right\}. \end{aligned}\right\} \tag{18.9}$$

Our task is to determine the mean velocity of diffusion of the kind of ion considered. We shall suppose that the ion density ϱ_1 and the temperature T depend on y only.

Let us consider ions which crossed an area dS in the plane $y = 0$ at a time between t' and $t' + dt'$ and have velocities between $\mathbf{v}_1$ and $\mathbf{v}_1 + d\mathbf{v}_1$ (Fig. 48). At time t' these ions were in a cylinder with base area dS and height $v\,dt'$. (If they had no collisions between t and t', their velocities at time t were between $\mathbf{v}_1'$ and $\mathbf{v}_1' + d\mathbf{v}_1'$.)

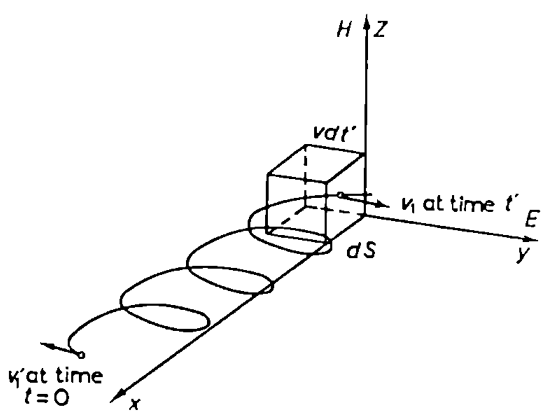

FIG. 48

The composition of the molecules which cross dS after a time $t' - t$ will vary owing to collisions. The number of collisions which ions of the kind considered undergo in 1 cm³ in 1 second is n_1/τ_1. Let $\chi_1(\mathbf{v}_1', y)\,d\mathbf{v}_1'$ be the fraction of collisions by which molecules with velocities between $\mathbf{v}_1'$ and $\mathbf{v}_1' + d\mathbf{v}_1'$ are formed, where the velocity is prescribed not only in magnitude, but in direction also. If the gas is in stable equilibrium, χ_1 is determined by Maxwell's function; in the general case, χ_1 differs very little from this if the drift velocity is small

compared with the velocity of thermal motion (which is true for the electric fields usually encountered). The number of collisions which give ions with velocities between v_1' and $\mathrm{v}_1' + \mathrm{d}\mathrm{v}_1'$ in a cylinder of volume $v\,\mathrm{d}t'\,\mathrm{d}S$ in the time from t to $t + \mathrm{d}t$ is

$$\chi_1(\mathrm{v}_1', y')\,\mathrm{d}\mathrm{v}_1'\,v\,\mathrm{d}S\,\mathrm{d}t'\,n_1\,\mathrm{d}t/\tau_1 .$$

A fraction $e^{-(t'-t)/\tau_1}$ of these particles reach $\mathrm{d}S$ without collisions. Thus the number of particles crossing $\mathrm{d}S$ during $\mathrm{d}t'$ with velocities between v_1 and $\mathrm{v}_1 + \mathrm{d}\mathrm{v}_1$ is

$$n_1\,v\,\mathrm{d}\mathrm{v}_1'\,\mathrm{d}S\,\mathrm{d}t' \int\limits_{-\infty}^{t'} \chi_1(\mathrm{v}_1', y')\,e^{-(t'-t)/\tau_1}\,\mathrm{d}t/\tau_1 .$$

In order to obtain the number of ions crossing $\mathrm{d}S$ with all velocities v_1, we must integrate the expression obtained over all v_1'. It is possible to change the variable of integration, since there is a one-to-one correspondence, defined by (18.8), between v_1 and v_1'. Here the cylinder in which collisions take place has a different position and volume for each value of v_1'. Changing the order of integration over the two independent variables, and replacing $t' - t$ by t, the time which elapses between the moment of collision and that of crossing $\mathrm{d}S$, we find for the number of ions of the given kind which cross $\mathrm{d}S$ in one direction in time $\mathrm{d}t'$ the expression

$$\mathrm{d}N = \mathrm{d}S\,\mathrm{d}t'\,n_1 \int\limits_0^{\infty} \left\{ \int \chi_1(\mathrm{v}_1', y')\,v\,\mathrm{d}\mathrm{v}_1' \right\} e^{-t/\tau_1}\,\mathrm{d}t/\tau_1 . \qquad (18.10)$$

This can be expressed in the form $\mathrm{d}N = n_1(0)\,\bar{v}_1\,\mathrm{d}S\,\mathrm{d}t'$, where $\bar{v}_1$ is the mean velocity of diffusion along the y axis.

Equating the right-hand sides and expanding $\chi_1(\mathrm{v}_1', y')$ in powers of y', we find

$$n_1(0)\,\bar{v}_1 = \int\limits_0^{\infty} \left\{ \int \left[n_1\,\chi_1(\mathrm{v}_1', 0) + y'\,\frac{\partial n_1\,\chi_1(\mathrm{v}_1', 0)}{\partial y} \right] v\,\mathrm{d}\mathrm{v}_1' \right\} e^{-t/\tau_1}\,\mathrm{d}t/\tau_1 , \quad (18.11)$$

where the higher powers of y' have been neglected, as is usual in diffusion theory. By means of equations (18.8) and (18.9) we can express v and y' in terms of u', v', x, y and t. For ease of writing we put the co-ordinates x and y of the area $\mathrm{d}S$ equal to zero, and we neglect terms in F_1^2, since, if the drift velocity is less than the thermal velocity, $m_1 F_1$ is less than the force defined by equation (18.1). We also drop the odd powers of u' and v', since the velocity distribution immediately after a collision does not differ greatly from the equilibrium distribution, i. e. it is almost isotropic, and these terms vanish on integrating over all velocities.

As a result we have

$$n_1(0)\,\bar{v}_1 = \int_0^\infty \Big[F_1 \int n_1\,\chi_1(\mathrm{v}_1',0)\,\mathrm{d}\mathrm{v}_1' - \frac{\partial}{\partial y}\int n_1\,\chi_1(\mathrm{v}_1',0)\,[v'^2\cos\omega_1 t + u'^2(1-\cos\omega_1 t)]\,\mathrm{d}\mathrm{v}_1'\Big\}\sin\omega_1 t\cdot e^{-t/\tau_1}\,\mathrm{d}t/\omega_1\tau_1\,. \tag{18.12}$$

Since the integration extends over all v_1', we can write

$$\int n_1\,\chi_1(\mathrm{v}_1',0)\,\mathrm{d}\mathrm{v}_1' = n_1\,.$$

Moreover, the mean value of u'^2 over all the particles is equal to that of v'^2, because of the isotropy of the velocity distribution after a collision. Hence

$$n_1\int\chi_1(\mathrm{v}_1',0)\,m_1\,u'^2\,\mathrm{d}\mathrm{v}_1' = n_1\int\chi_1(\mathrm{v}_1',0)\,m_1\,v'^2\,\mathrm{d}\mathrm{v}_1' = p_1\,,$$

where p_1 is the partial pressure of the ions of the kind considered on the plane $y = 0$.

Substituting the equations so obtained in (18.12), we find, using the equality of the mean values of v'^2 and u'^2, and $n_1(0)\,m_1 = \varrho_1$,

$$\begin{aligned}\bar{v}_1 &= \left(F_1 - \frac{1}{\varrho_1}\frac{\partial p_1}{\partial y}\right)\int_0^\infty \sin\omega_1 t\cdot e^{-t/\tau_1}\,\mathrm{d}t/\omega_1\tau_1\\ &= \left(F_1 - \frac{1}{\varrho_1}\frac{\partial p_1}{\partial y}\right)\frac{\tau_1}{1+\omega_1^2\tau_1^2}\,.\end{aligned} \tag{18.13}$$

For $H = 0$ $(\omega_1 = 0)$, the mean velocity of diffusion would be

$$\bar{v}_1 = \tau_1\left(F_1 - \frac{1}{\varrho_1}\frac{\partial p_1}{\partial y}\right).$$

Thus the presence of the magnetic field diminishes the velocity of diffusion by a factor $1 + \omega_1^2\tau_1^2$.

Besides the *direct* diffusion in the direction of F_1, a *transverse* diffusion takes place in a magnetic field, perpendicular to both fields and along the x axis. We find, by a method similar to the foregoing

$$n_1(0)\,\bar{u}_1 = \int_0^\infty \Big\{\int\Big[n_1\,\chi_1(\mathrm{v}_1',0) + y'\frac{\partial n_1\chi_1(\mathrm{v}_1',0)}{\partial y}\Big]\,u\,\mathrm{d}\mathrm{v}_1'\Big\}\,e^{-t/\tau_1}\,\mathrm{d}t/\tau_1\,. \tag{18.14}$$

Replacing y' and u by means of (18.8) and (18.9) and making the same assumptions, we find

$$\bar{u}_1 = \left(F_1 - \frac{1}{\varrho_1}\frac{\partial p_1}{\partial y}\right)\int_0^\infty (1-\cos\omega_1 t)\, e^{-t/\tau_1}\,\mathrm{d}t\,\omega_1\tau_1$$
$$= \left(F_1 - \frac{1}{\varrho_1}\frac{\partial p_1}{\partial y}\right)\frac{\omega_1\tau_1^2}{1+\omega_1^2\tau_1^2}\,. \tag{18.15}$$

We see that the mean velocity of transverse diffusion is $\omega_1\tau_1$ times the mean velocity of direct diffusion.

Similar expressions are obtained for the other kinds of ions and for electrons. The current density is the sum of the current densities formed by the motion of electrons and ions. We obtain for the *ionic* component of the current density, replacing F_1 by the expression for it, the formula

$$\mathbf{j}_i = n_i e\bar{\mathbf{v}}_i = \frac{n_i e^2}{m_i}\frac{\tau_i}{1+\omega_i^2\tau_i^2}\left(\mathbf{E} - \frac{1}{n_i e}\frac{\mathrm{d}p_i}{\mathrm{d}\mathbf{r}}\right) +$$
$$+ \frac{n_i e^2}{m_i}\frac{\omega_i\tau_i^2}{1+\omega_i^2\tau_i^2}\cdot\frac{1}{H}\left(\mathbf{E} - \frac{1}{n_i e}\frac{\mathrm{d}p_i}{\mathrm{d}\mathbf{r}}\right)\wedge\mathbf{H}\,. \tag{18.16}$$

We can introduce the direct and transverse conductivities for ions, $\lambda_{i,\mathrm{I}}$ and $\lambda_{i,\mathrm{II}}$, and also for electrons. If $\omega_i\tau_i \gg 1$, the total transverse current is easily found to be zero, and the direct ionic conductivity is $\surd(m_i/m_e)$ times the direct electron conductivity. The opposite is true in the absence of a magnetic field. It should be noticed that we have carried out the calculation in a co-ordinate system fixed to the centre of mass of the gas, since we have supposed the velocity distribution after a collision to be isotropic.

3. The interaction of electrons and ions. In some layers of the Sun's atmosphere the ionisation is considerable, so that we cannot neglect the collisions between ions and electrons. In order to discuss this case in greater detail, we must first study the mechanism of the mutual interaction of ions and electrons.

Let an ion of mass m_1 and charge Ze be stationary, while another ion of mass m_2 and charge e moves past it with velocity $\mathbf{v}$ and closest approach distance p (the distance from the stationary ion to the straight line in which the second ion would move if the ions did not interact).

In the co-ordinate system in which the centre of mass of the two particles is at rest, the first ("stationary") ion moves with velocity

$-m_2 \mathbf{v}/(m_1+m_2)$, and the second ("colliding") ion with velocity $m_1 \mathbf{v}/(m_1+m_2)$. We know from mechanics that in this co-ordinate system the two ions move in hyperbolae, the centre of mass of the particles being at the external focus (if the particles repel each other) or the internal focus (if they attract each other). The magnitude of the velocity of either ion is the same long before and long after the collision (the asymptotic velocity), but the direction of the velocity is changed, as a result of the interaction, by a certain angle (Fig. 49).

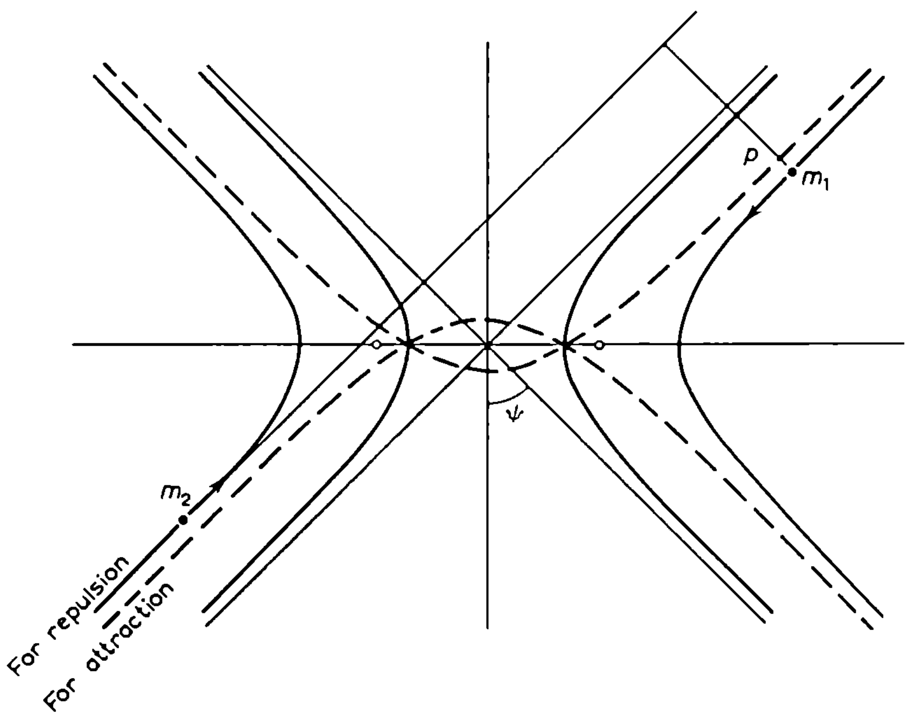

Fig. 49

After the collision, the component velocity of the first ion in the direction of its original velocity is

$$v_1 = -\frac{m_2}{m_1+m_2} v \cos 2\psi , \tag{18.17}$$

and the component perpendicular to this is

$$v_2 = \frac{m_2}{m_1+m_2} v \sin 2\psi . \tag{18.18}$$

In the co-ordinate system which was originally stationary with respect to the first ion m_1, these velocity components are respectively

$$\begin{aligned} v_1' &= \frac{m_2}{m_1+m_2} v - \frac{m_2}{m_1+m_2} v \cos 2\psi = \frac{m_2}{m_1+m_2} v(1-\cos 2\psi)\,, \\ v_2' &= \frac{m_2}{m_2+m_1} v \sin 2\psi\,. \end{aligned} \tag{18.19}$$

The solution of the equations of motion of the ion gives for ψ (half the angle between the asymptotes) the expression

$$\cot\psi = \pm \frac{m_1 m_2}{m_1+m_2} \frac{v^2 p}{Z e^2} = \pm \frac{p}{a}\,, \tag{18.20}$$

where

$$a = \frac{Z e^2}{v^2} \frac{m_1+m_2}{m_1 m_2}\,; \tag{18.21}$$

a is the distance of closest approach for which the direction of motion of the particle is turned through an angle $\frac{1}{2}\pi$.

Substituting for ψ in (18.19), we find

$$\begin{aligned} v_1' = 2 \frac{m_2 v}{m_1+m_2} \cos^2\psi &= \frac{2 m_2 v}{m_1+m_2} \frac{1}{1+p^2/a^2} \\ &= \frac{2 Z^2 e^4}{v^3} \frac{m_1+m_2}{m_1^2 m_2} \cdot \frac{1}{p^2+a^2}\,. \end{aligned} \tag{18.22}$$

If the number of m_2 ions in 1 cm^3 is n, and they all move with velocity v, then $\mathrm{d}n = n\,v \cdot 2\pi\,p\,\mathrm{d}p$ is the number of m_2 ions which pass the m_1 ion in 1 second at distances between p and $p + \mathrm{d}p$. Each of these gives the m_1 ion a velocity with components v_1' and v_2'. The contributions v_2' from different ions cancel one another, and the total change in velocity is

$$\mathrm{d}v = n \cdot 2\pi \frac{2 Z^2 e^4}{v^2} \cdot \frac{m_1+m_2}{m_1^2 m_2} \frac{p\,\mathrm{d}p}{p^2+a^2}\,. \tag{18.23}$$

In order to find the acceleration of the ion, i. e. the total change in velocity in 1 second, we must integrate dv over all possible values of p. However, as $p \to \infty$ the integral diverges. Hence we must restrict the value of p. We denote the upper limit of the values of p by D. Then the force acting on m_1 is

$$m_1 \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{2\pi Z^2 e^4}{v^2} \cdot \frac{m_1+m_2}{m_1 m_2} n \log_e \left(1 + \frac{D^2}{a^2}\right)\,, \tag{18.24}$$

where, as we have said,

$$a = Z e^2 (m_1+m_2)/m_1 m_2 v^2\,.$$

The problem which we have considered is equivalent to that of the retardation of an ion m_1 moving in a plasma with a velocity v which is considerably greater than the thermal velocity of the ions in the plasma (we regard the latter as stationary). If an ion (or electron) of mass m_1 is moving with a velocity greater than that of the electrons forming the plasma [m_2 in formulae (18.17) to (18.24) is the mass of an electron], then the retardation will be produced chiefly by the electrons, whose retarding force is greater, and the retarding force of the ions can be neglected. Various values are taken for D in the work of different authors. Sometimes it is put equal to what is called the Debye radius $D = \sqrt{(kT/8\pi\, n\, e^2)}$, the distance at which the field of the faster ion is screened by the electrostatic polarisation of the plasma*. More often, however, D is put equal to half the mean distance between the ions $(\frac{1}{2}\, n^{-\frac{1}{3}})$, so that not more than one particle is in the sphere of interaction at one time, because we have assumed in deriving (18.24) that the collisions are between pairs of particles.

The many-particle theory takes account of further interactions which are not between pairs. These differences, however, have very little effect on the value of the retarding force, since the quantity D appears only in a logarithm. We shall take $D = \frac{1}{2}\, n^{-\frac{1}{3}}$. In ordinary conditions $D \gg a$, and the retarding force is

$$m_1 \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{4\pi Z^2 e^4}{v^2} \cdot \frac{m_1 + m_2}{m_1 m_2}\, n \log_e \left(\frac{m_1 m_2}{m_1 + m_2} \cdot \frac{v^2}{2 Z e^2 n^{\frac{1}{3}}}\right). \qquad (18.25)$$

If the ion is moving with the thermal velocity, the force which retards its motion, in the direction of motion, is of the same order of magnitude, but the mean retardation is almost zero, since it not only loses energy, but also gains it from other ions. The retarding force on an electron in the plasma, due to other electrons, is

$$m_e \frac{\mathrm{d}v}{\mathrm{d}t} \approx \frac{8\pi e^4 n_e}{3kT} \log_e \frac{3kT}{4 e^2 n_e^{\frac{1}{3}}} = \frac{8\pi e^4 n_e}{3kT} L, \qquad (18.26)$$

where $L = \log_e(3kT/4e^2 n_e^{\frac{1}{3}})$. In the various layers of the Sun's atmosphere, L varies from 10 to 20. From this we find the mean free path

$$l = \frac{v}{\mathrm{d}v/\mathrm{d}r} = \frac{v^2}{\mathrm{d}v/\mathrm{d}t} \approx \frac{(kT)^2}{3 e^4 n_e L}, \qquad (18.27)$$

and the effective cross-section

$$\sigma \approx 1/l\, n_e = 3\, e^4 L/(kT)^2. \qquad (18.28)$$

We see that l and σ are independent of the mass of the ion.

* For the slower ion, $D = \sqrt{(kT/4\pi n e^2)}$.

We notice that, if we considered only strong interactions which deflect the particle through an angle greater than 90°, the closest approach would be defined by the condition of equality of the potential and kinetic energies of the particle $e^2/p = 3\,k\,T/2$, whence the cross-section for close collisions $\sigma' = \pi\,p^2 = 4\,\pi\,e^4/9\,(k\,T)^2$. Thus taking account of distant interactions increases the effective cross-section by a factor of approximately $2\,L$.

The interaction of electrons with electrons and of ions with ions, which we have considered, determines the velocity distribution function of the ions and electrons separately. This interaction, however, does not alter the mean velocity of the electron gas or ion gas as a whole, i. e. it has no effect on the value of the total conductivity of the plasma.

This conductivity is determined by the interaction of electrons with ions, i. e. by the time of free flight of an electron, which is limited by the retardation due to the ions, and conversely. Using formula (18.25), we find that l_e determined by ions is equal to $2\,l$. Then the mean time of free flight

$$\tau_e = \frac{l_e}{\bar{v}_e} = \frac{(k\,T)^2\,m_e^{\frac{1}{2}}}{3\,e^4\,n_e\,L\,(3\,k\,T)^{\frac{1}{2}}} = \frac{(k\,T)^{3/2}\,m_e^{\frac{1}{2}}}{3\sqrt{3}\,e^4\,n_i\,L} \text{ and } \tau_i = \frac{(k\,T)^{3/2}\,m_i^{\frac{1}{2}}}{3\sqrt{3}\,e^4\,n_e\,L}, \tag{18.29}$$

while the time of free flight of an individual particle is now proportional to the cube of its velocity. Furthermore, another complication appears: we cannot neglect the exchange of momenta between electrons and ions. To determine the current density we can no longer, in general, use the elementary theory, but must start from the kinetic equation. The corresponding calculations are very complicated, and we shall not go through them, but give only the final result (due to T. G. COWLING [34]).

The velocities of diffusion of the individual components are not determined by the equation: to find them, it is necessary to find independently the motion of the gas as a whole. Only the *relative velocity of diffusion* under the action of forces perpendicular to the magnetic field is determined in the first approximation; it is

$$\bar{\mathbf{v}}_i - \bar{\mathbf{v}}_e = \frac{\tau\left(\mathbf{F}_i - \frac{1}{\varrho_i}\frac{\mathrm{d}p_i}{\mathrm{d}\mathbf{r}} - \mathbf{F}_e + \frac{1}{\varrho_e}\frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}}\right)}{1+\omega^2\tau^2} + \frac{\omega\,\tau^2\,H^{-1}\left(\mathbf{F}_i - \frac{1}{\varrho_i}\frac{\mathrm{d}p_i}{\mathrm{d}\mathbf{r}} - \mathbf{F}_e + \frac{1}{\varrho_e}\frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}}\right)_\wedge \mathbf{H}}{1+\omega^2\tau^2}, \tag{18.30}$$

where

$$\tau = \frac{\tau_i\,\varrho_e + \tau_e\,\varrho_i}{\varrho_e + \varrho_i} \approx \tau_e\,,\ \omega = \frac{\omega_i\,\varrho_e + \omega_e\,\varrho_i}{\varrho_e + \varrho_i} \approx \omega_e = -\frac{e\,H}{m\,c}\,.$$

At first sight the expressions (18.13) and (18.15) appear similar to the expression (18.30), but there is an important difference between them. In the former, the velocity of diffusion of each gas depends only on the forces acting on that gas, and on ω and τ for that gas, while in the latter case the velocity of diffusion of each gas depends on the forces acting on all the gases, and on the weighted mean values of ω and τ.

For the current density in the second case we obtain, on neglecting small terms,

$$\mathbf{j} = \frac{n_e e^2}{m} \left\{ \frac{\tau}{1 + \omega^2 \tau^2} \left(\mathbf{E} + \frac{1}{e n_e} \frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}} \right) + \frac{\omega \tau^2}{1 + \omega^2 \tau^2} \cdot \frac{1}{H} \left(\mathbf{E} + \frac{1}{e n_e} \frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}} \right)_{\wedge} \mathbf{H} \right\}$$

$$= \lambda_{\mathrm{I}} \left(\mathbf{E} + \frac{1}{e n_e} \frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}} \right) + \lambda_{\mathrm{II}} H^{-1} \left(\mathbf{E} + \frac{1}{e n_e} \frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}} \right)_{\wedge} \mathbf{H}\,, \qquad (18.31)$$

where $\lambda_{\mathrm{I}} = \lambda_0/(1 + \omega^2\tau^2)$ and $\lambda_{\mathrm{II}} = \lambda_0\, \omega\, \tau/(1 + \omega^2\tau^2)$ are the direct and transverse conductivities. If $\omega\, \tau \gg 1$, which holds for the chromosphere when $H \geq 10^{-2}$, and in the corona when $H \geq 10^{-5}$, we have $\lambda_{\mathrm{I}} \ll \lambda_{\mathrm{II}} \ll \lambda_0$.

It remains to decide which of the formulae for the current density should be used for the solar atmosphere. For $T = 5000°$, $\sigma \approx 10^{-12}$, whereas a neutral atom has a cross-section $\sigma_0 \approx 10^{-15}$. This means that, if the number of ions is more than one thousandth of the number of neutral atoms, formula (18.31) must be used. Formula (18.16) may be used only in sunspots and possibly in a thin layer of the photosphere where the ionisation is least. Formula (18.31) is widely used in theories of the electromagnetic phenomena in the Sun's atmosphere.

4. Electric fields on the Sun. If there is a pressure gradient in the plasma, it produces a current independently of the presence of an electric field (18.31). This current results in the formation of a volume charge whose field opposes the further separation of charges. The value of the field at which the current ceases is, by (18.31),

$$\mathbf{E} = -\frac{1}{e n_e} \operatorname{\mathbf{grad}} p_e\,. \qquad (18.32)$$

This field, as we see from the expression for it, is a conservative field if there is a one-to-one relation between p_e and n_e. The amount of charge which produces it is very small; it is easily found that the number of unbalanced elementary charges is many orders of magnitude less than the number of ions in the same volume. The density decrease with height in the atmosphere produces in the photosphere a field $E_0 \approx 5 \times 10^{-11}$ abs. esu $\approx 1{\cdot}5 \times 10^{-8}$ volt/cm.

If an isothermal atmosphere of ionised hydrogen is in hydrostatic equilibrium, the equilibrium conditions for electrons and for protons can be written separately in the form

$$\begin{aligned} \mathrm{d}p_e/\mathrm{d}\mathbf{r} &= n_e\, m_e\, \mathbf{g} + n_e\, e\mathbf{E}\,, \\ \mathrm{d}p_i/\mathrm{d}\mathbf{r} &= n_i\, m_i\, \mathbf{g} - n_i\, e\mathbf{E}\,, \end{aligned} \tag{18.33}$$

where $n_e = n_i$ and $p_e = p_i$, and the intensity vector points upwards. We find from the equations (18.33)

$$\frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}} = \tfrac{1}{2}\, m_{\mathrm{H}}\, n_e\, \mathbf{g}, \quad \mathbf{E} = -\frac{1}{e} \cdot \tfrac{1}{2}\, m_{\mathrm{H}}\, \mathbf{g} = -\frac{1}{e n_e} \frac{\mathrm{d}p_e}{\mathrm{d}\mathbf{r}}\,. \tag{18.34}$$

We see that the field exerts on the proton a force $e\mathbf{E}$ equal to half its weight, and the effective weight of the proton, i. e. the force of attraction on it towards the Sun, is now equal to the "effective weight" of the electron $-e\mathbf{E}$; this brings about the same height distribution of both protons and electrons, as is necessary for the atmosphere to be macroscopically neutral.

Other than the above, we know of hardly any mechanisms which lead to the formation of a volume charge in an ionised gas. Its large conductivity causes a dissipation of the charge in a time of the order of λ^{-1}, i. e. in 10^{-12} second, and hence considerable potential fields are impossible. *Electric eddy fields* are much more important, since they do not lead to the formation of volume charges. The variation of sunspots is the chief source of eddy fields. Any change in the magnetic flux causes the appearance of an induction field determined by the condition

$$\oint_L E_s\, \mathrm{d}s = -\frac{1}{c} \frac{\partial \Phi}{\partial t}\,, \tag{18.35}$$

where E_s is the projection of $\mathbf{E}$ on the contour L, and Φ is the magnetic induction flux through the surface bounded by L.

If a spot with field intensity $H = 10^3$ abs. emu and radius $r = 5 \times 10^8$ cm appears in a time $t = 3$ days, the maximum intensity of the electric eddy field above the spot is given by the expression

$$\begin{aligned} E = \frac{1}{2\pi r} \frac{1}{c} \frac{\pi r^2 H}{t} = \frac{r}{2c} \frac{H}{t} &= \frac{5 \times 10^8 \,.\, 10^3}{2 \times 3 \times 10^{10} \,.\, 2{\cdot}4 \times 10^5} \\ &\approx 3 \times 10^{-5} \text{ abs. esu} = 10^{-2} \text{ volt/cm.} \end{aligned}$$

However, this field is referred to a stationary co-ordinate system. If we wish to consider the action of the field in some medium, we must regard the contour L as fixed to that medium (L is not a geometrical

contour, but a material one), and we must take account of the motion of the medium in calculating the field in a co-ordinate system fixed to the plasma. If, for example, the matter above the growing spot is compressed at such a rate that the flux through the compressed contour is unchanged, there will be no induction currents, and the field in the moving medium will be zero.

Let some body be moving in the field $\mathbf{H}$ with velocity $\mathbf{v}$. If a charge e is attached to this body, it is acted on during the motion by a force

$$\mathbf{f} = e\,\frac{1}{c}\,\mathbf{v} \wedge \mathbf{H} = e\mathbf{E}', \tag{18.36}$$

where

$$\mathbf{E}' = \frac{1}{c}\,\mathbf{v} \wedge \mathbf{H}\,.$$

An observer who is stationary relative to the body might suppose that in his system of reference there is an electric field $\mathbf{E}'$. This phenomenon is a particular case of electromagnetic induction, which is very important in the electrodynamics of the Sun. Let us consider an example which illustrates this. Let an extremely rarefied plasma be in fields $\mathbf{H}$ and $\mathbf{E}$. As we said above, the plasma will drift with a velocity $\mathbf{v} = c\mathbf{E} \wedge \mathbf{H}/H^2$. This motion produces a field $\mathbf{E}' = (1/c)\,\mathbf{v} \wedge \mathbf{H} = -\mathbf{E}$. Thus the resultant field in the moving plasma is $\mathbf{E} + \mathbf{E}' = 0$, and no currents arise in it.

The interaction of the particles diminishes the drift velocity; in this case also, however, we must substitute in the formula for the current density not the value of $\mathbf{E}$ in the system stationary relative to the observer, but the value of $\mathbf{E}$ in the system of reference fixed to the plasma as a whole. In other words, in order to determine $\mathbf{j}$, which is independent of the choice of co-ordinate system, we have to use the value of $\mathbf{E}$ in a definite system of reference.

If the conducting medium moves in a magnetic field under the action of some forces, then $\mathbf{E}'$, like any other field, causes the appearance of a current given by (18.31).

Besides the cases we have considered, induction fields appear when the magnetic moment of a bipolar spot group is changed in various ways: when the distance between the spots changes, when the group rotates about its centre of gravity, and so on. The latter case is of especial interest, since here the electric field is parallel to the magnetic field, and the conductivity is high.

The appearance of currents produces a new electric field, the *self-induction field* $\mathbf{E}''$, which is of great importance under the conditions of high conductivity and large dimensions found in the solar atmosphere. This field arises as a result of the changing magnetic field of the currents. The current density is determined by the combined effect of all the electric fields, $\mathbf{E} + \mathbf{E}' + \mathbf{E}''$.

5. **Thermal and mechanical effects of currents.** Currents which arise as different manifestations of electromagnetic induction can produce thermal and mechanical effects. The heat evolved by a current in 1 cm³ in 1 second is determined by the expression

$$q = \mathbf{j} \cdot \mathbf{E} = \lambda E^2 = j^2/\lambda\,. \tag{18.37}$$

The amount of heat evolved is determined by the component of the current parallel to E, and so, in the case of an anisotropic medium, we must use λ_1 instead of λ.

It is possible that the cause of the high temperature of the corona and upper chromosphere, and of the appearance of faculae and flocculi, is the heating of these regions by electric currents. It is significant that spots are always surrounded by faculae and the temperature of the corona above them is higher.

The electric field communicates kinetic energy to the electrons. In ordinary conditions, when the fields are not very strong, the electrons pass on this energy to the ions, and the kinetic temperatures of ions and electrons are the same. However, an electron cannot pass on to the ion, at each elastic collision, more than the fraction $(4\, m_e/m_i)$ of its energy. We can find a value of the field $\mathbf{E} = \mathbf{E}_0$ such that the additional energy acquired by an electron over the distance of its free path is equal to the maximum fraction of thermal energy which it can pass on to an ion. The length of the free path is proportional to the square of the energy, and hence, if $E > E_0$, the electron will gain more energy than it can give up; the energy increases, the free path increases, the energy increases further, and so on. The energy of the electron can be limited only by inelastic collisions or by the finite dimensions of the field. Thus the field may bring about a discharge. This critical value of the field at which a discharge can take place is, according to R. G. GIOVANELLI,

$$E_0 = \frac{2\sqrt{3}}{9}\,\frac{\pi \beta^2 m_e^{3/2} n_i Z^2}{u_0^2\, e\, m_i^{\frac{1}{2}}}\,, \tag{18.38}$$

where $\beta = 1{\cdot}57 \times 10^9$ cm³/sec², and u_0 is the mean velocity of an electron in thermodynamic equilibrium with the ions. At a temperature $T = 5750°$, $E_0 = 7 \times 10^{-5}\, p_i$. In the chromosphere $p_i \approx 1$ dyne/cm², so that if $E \geq 10^{-4}$ a discharge can take place.

It must be borne in mind that a steady conservative field, arising, for example, from a concentration gradient, cannot be the cause of an evolution of heat, since it brings about only some redistribution of the electron density, and not a motion of the electrons. Formula (18.31) shows that the current density in this case is zero. Besides the eddy

fields, a current may be produced by a variable conservative field, but this is practically impossible in an ionised atmosphere.

The mechanical effect of a current is found from (18.2). The total force exerted on the gas is equal to the sum of the forces acting on the ions and on the electrons:

$$\begin{aligned} \mathbf{f} = \mathbf{f}_i + \mathbf{f}_e &= n_i e\mathbf{E} + \frac{n_i e}{c} \mathbf{v}_i \wedge \mathbf{H} - n_e e\mathbf{E} - \frac{n_e e}{c} \mathbf{v}_e \wedge \mathbf{H} \\ &= \frac{1}{c} n_e e (\mathbf{v}_i - \mathbf{v}_e) \wedge \mathbf{H} = \frac{1}{c} \mathbf{j} \wedge \mathbf{H}, \end{aligned} \tag{18.39}$$

since $n_i = n_e$, and $\mathbf{j}$ in (18.39) denotes the total current, both direct and transverse. The force $\mathbf{f}$ sets the plasma in motion, and this motion in turn alters the current. Hence the current cannot be calculated simply from the conductivity; it is necessary to solve simultaneously the conduction equations and the equations of motion of the plasma under the action of the forces $\mathbf{f}$, **grad** p, etc., and to take account of drift and of the self-induction determined by Maxwell's equations. If **grad** $p = 0$, the plasma moves and the current will be much less than in a stationary medium. It is necessary for equilibrium that the equation **grad** $p = \mathbf{j} \wedge \mathbf{H}/c$ should be satisfied. The pressure gradient, in turn, produces the current.

Thus the solution of the problem of finding the currents and the motion of matter in electromagnetic fields is very complex. The mechanical effect of the current on the gas is as yet the only known cause which could explain the motions of prominences. The electrostatic force acting on charged clouds is much less important, because of the small amount of volume charge possible. However, no theory has yet been constructed which satisfactorily explains the motions of prominences.

If a conducting medium moves in a magnetic field with a velocity $\mathbf{v}$ which is perpendicular to $\mathbf{H}$, a current density $\mathbf{j} = \lambda \mathbf{v} \wedge \mathbf{H}/c$ appears in it. In practice, the current due to self-induction will increase gradually, and may be several orders of magnitude less, depending on the actual conditions. This current exerts on 1 cm³ of the medium a force

$$\mathbf{f} = \frac{1}{c} \mathbf{j} \wedge \mathbf{H} = -\frac{\lambda H^2}{c^2} \mathbf{v}. \tag{18.40}$$

The force $\mathbf{f}$, as is easily seen, stops the motion of the medium in a time of the order

$$t \approx \varrho v/f = \varrho c^2/\lambda H^2, \tag{18.41}$$

or several orders of magnitude more because of self-induction, the effect of which we shall estimate later. The motion will not, in fact, be *uniformly retarded*; the appearance of induction fields, as we shall

show below, leads to a *wave motion*, or to an even more complicated motion. The example given shows only that steady uniform motion of the conducting medium across $\mathbf{H}$ is impossible in the absence of subsidiary forces.

Hence, if a change in the magnetic field, such as an intensification or concentration of the lines of force, takes place in the easily mobile atmosphere of the Sun, then the matter, which moves with the lines of force, will also be concentrated, forming a *condensation*, a denser region. The appearance of prominences above spots may be related to this circumstance. The direction of motion of prominences and currents may be determined by the direction of the field which retards the transverse motions.

Since the plasma follows the field, the magnetic flux through any material contour is invariable, and no inductive electromotive force arises. For the appearance of an induction current it is necessary that some force, such as a pressure gradient, should hinder the motion of the matter and prevent it from following the field. This may happen, for example, in the formation of a condensation above a growing spot. The same can occur if the motion perpendicular to the direction of the field is maintained by some force.

Let us consider this case in more detail. Let there be a force of density $\mathbf{f}'$ supporting the motion. *In steady motion*, $\mathbf{f}'$ will be balanced by a force $\mathbf{f} = \mathbf{j} \wedge \mathbf{H}/c$. Thus the magnitude of the force $\mathbf{f}'$ determines the current density (for a given $\mathbf{H}$). Self-induction is unimportant here, since the current density is constant in the steady state. The velocity of motion of the matter relative to the field depends on the conductivity: the greater the conductivity, the less the velocity which produces the necessary current density. In ordinary conditions this velocity is small.

Let us calculate, for example, the rate of fall of an ionised prominence in a magnetic field whose intensity is directed horizontally and is $H = 1$ gauss, while $n_e = 10^{10}$. We obtain from (18.40) the condition for a steady state: $n_e m_{\mathrm{H}} g = \lambda_{\mathrm{I}} H^2 v/c^2$, where $g = 2{\cdot}7 \times 10^4$ is the acceleration due to gravity. According to (18.29), putting $T = 5000°$, we find $\tau = \tau_e = 10^{-6}$, $\omega = \omega_e = -2 \times 10^7$ and $|\omega\tau| = 20 \gg 1$. In this case

$$\lambda_{\mathrm{I}} = \frac{\lambda_0}{\omega^2\tau^2} = \frac{n_e e^2 \tau}{m_e} \frac{1}{\tau^2} \cdot \frac{m_e^2 c^2}{e^2 H^2} = \frac{n_e m_e c^2}{\tau H^2},$$

and the rate of fall

$$v = n_e m_{\mathrm{H}} g c^2/\lambda_{\mathrm{I}} H^2 = m_{\mathrm{H}} g \tau/m_e \approx 20 \text{ cm/sec}.$$

The order of magnitude is not changed when drift is taken into account. In practice, this means that the prominence will simply hang above the Sun's surface. The existence of stationary prominences may be

related to the presence of strictly horizontal fields or fields whose lines of force contain a cavity. If, however, the field intensity vector is inclined to the Sun's surface, the prominence will move in the direction of the field intensity, i. e. in an inclined plane, with the appropriate acceleration.

Besides the case we have considered, another in which the matter moves relative to the field is in principle possible. This occurs when a large mass of gas moves from a region where the intensity is low to one where it is high. As we shall show below, a field cannot at once penetrate into a cloud of gas which conducts a current; screening currents are induced on the surface of the cloud, and their magnetic field neutralises the field inside the cloud. Thus the interior of the cloud is not subjected to a retarding force, and the retardation time is increased.

The interaction of the screening currents with the magnetic field produces what is called a magnetic pressure, equal to $H^2/8\pi$. We shall estimate its retarding effect. Let there be a cloud of length $l = 1000$ km and density $\varrho = 10^{-12}$ g/cm^3, moving with velocity $v = 30$ km/sec in a field $H = 30$ gauss. Then the motion is stopped in a time $t \approx l \varrho v \,.\, 8\pi/H^2 = 8$ seconds. We see that "electromagnetic screening" can be of importance only in the case of very large masses and weak fields. On the whole, we can draw the conclusion that the motion of an ionised gas across the magnetic field is stopped comparatively rapidly if there is no considerable force supporting the motion.

6. Allowance for self-induction. We have assumed in our calculations that the current density has a steady value $\mathbf{j} = \lambda \mathbf{E}$. However, it is known that, when a field $\mathbf{E}$ is applied, the current does not immediately reach its maximum value; it rises gradually, and the time for it to rise to maximum is proportional to the self-induction of the conductor. The self-induction of a large volume of conducting material is large; the current density increases slowly, and the retarding force also increases slowly, so that the retardation time is increased. To obtain a quantitative result, we must simultaneously solve Maxwell's equations and the hydrodynamical equation. This solution has been effected by H. ALFVÉN for some particular cases. He starts from the equations

$$\operatorname{curl} \mathbf{H} = \frac{4\pi}{c} \mathbf{j}\,, \tag{18.42}$$

$$\operatorname{curl} \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}\,, \tag{18.43}$$

$$\mathbf{j} = \lambda\left(\mathbf{E} + \frac{1}{c}\mathbf{v}_\wedge \mathbf{H}\right), \tag{18.44}$$

$$\varrho\frac{d\mathbf{v}}{dt} = \frac{1}{c}\mathbf{j}_\wedge \mathbf{H} - \operatorname{grad} p\,, \tag{18.45}$$

where $\mathbf{E}$ is the electric self-induction field.

In formulae (18.42) to (18.45) we have used the results that the displacement current in the solar atmosphere is considerably less than the conduction current, and the magnetic permeability is unity.

As a first approximation, let us consider a medium of constant density with an infinite conductivity, the velocity of the medium at the initial instant being perpendicular to the uniform magnetic field $\mathbf{H}_0$, and its magnitude depending only on z (the z axis is parallel to $\mathbf{H}$). Then the magnitudes of all the vectors appearing in the equations depend on z and t only. We rotate the co-ordinate system so that $j_y = 0$. Further, $j_z = 0$, since the induction fields in this case are perpendicular to $\mathbf{H}$. Then we find from (18.42)

$$j_x = -\frac{c}{4\pi}\frac{\partial H_y}{\partial z}, \quad H_x = \text{constant} = 0, \quad H_z = H_0 . \tag{18.46}$$

A field H_y due to the induced current system is added to the original magnetic field H_0.

We substitute these values in (18.45). Since, by hypothesis, **grad** p has no component in the xy plane, we obtain

$$\left.\begin{aligned} &\frac{\partial v_x}{\partial t} = 0, \quad v_x = \text{constant} = 0 , \\ &\frac{\partial v_y}{\partial t} = \frac{H_0}{4\pi\varrho}\frac{\partial H_y}{\partial z}, \quad v_z = 0 \end{aligned}\right\} \tag{18.47}$$

and also

$$\frac{\partial p}{\partial z} = -\frac{1}{8\pi}\frac{\partial (H_y^2)}{\partial z} . \tag{18.48}$$

Since $\lambda = \infty$, while $\mathbf{j}$ is finite, equation (18.44) gives

$$\mathbf{E} = -\frac{1}{c}\mathbf{v}_\wedge \mathbf{H}, \quad E_x = -\frac{1}{c} v_y H_0 , \quad E_y = E_z = 0 . \tag{18.49}$$

We find from equation (18.43)

$$\partial H_y/\partial t = - c\, \partial E_x/\partial z . \tag{18.50}$$

Using (18.49), (18.50) and (18.47), we obtain

$$\frac{\partial^2 H_y}{\partial t^2} = \frac{H_0^2}{4\pi\varrho}\frac{\partial^2 H_y}{\partial z^2} . \tag{18.51}$$

The equation we have obtained is identical with the wave equation, and consequently it describes a wave motion propagated along the z axis with a velocity

$$V = H_0/\surd(4\pi\varrho) . \tag{18.52}$$

The state of motion of the fluid moves along $\mathbf{H}_0$ with a velocity proportional to the field intensity and inversely proportional to $\sqrt{\varrho}$.

These magneto-hydrodynamic waves play a very great part in the physics of the Sun: they are capable of transmitting motion and variable magnetic fields from one layer to another. If the wave is sinusoidal, the amplitude of oscillation of H_y being A and the frequency Ω, then equations (18.46) to (18.50) give the following expressions for all the other quantities:

$$\left.\begin{aligned} H_y &= A \sin \Omega\,(t - z/v)\,, \\ v_y &= -\frac{A}{\sqrt{(4\pi\varrho)}} \sin \Omega \left(t - \frac{z}{v}\right), \\ j_x &= A\,\frac{c\Omega}{H} \sqrt{\left(\frac{\varrho}{4\pi}\right)} \cos \Omega \left(t - \frac{z}{v}\right), \\ E_x &= A\,\frac{H_0}{c\sqrt{(4\pi\varrho)}} \sin \Omega \left(t - \frac{z}{v}\right), \\ p &= p_0 - \frac{A^2}{8\pi} \sin^2 \Omega \left(t - \frac{z}{v}\right). \end{aligned}\right\} \tag{18.53}$$

The magnetic lines of force, which were previously straight lines, are now transformed into sinusoids by the superposition of the field H_y. Their form is determined by the equation

$$\mathrm{d}y/\mathrm{d}z = H_y/H_z\,, \tag{18.54}$$

whose solution is

$$x = x_0,\ y = y_0 + \frac{A}{\Omega\sqrt{(4\pi\varrho)}} \cos \Omega \left(t - \frac{z}{v}\right). \tag{18.55}$$

The velocity of the motion of the lines of force is $\mathrm{d}y/\mathrm{d}t$, which is easily seen to be equal to v_y.

The concepts of *mechanical* motion are, generally speaking, inapplicable to the field and to the lines of force. We introduce these concepts as an easily grasped analogy, which often allows the nature of the phenomenon as a whole to be qualitatively imagined. The results obtained by the use of such a method cannot be considered to be rigorously justified.

Thus the matter and the lines of force move in unison. This is to be expected from the general properties of a conducting medium "frozen" to the lines of force; see below. From this standpoint, we can regard the magneto-hydrodynamic waves as waves running on a stretched elastic string which corresponds to a line of force, its mass being $m = \varrho/H_0$ per centimetre, since the number of lines of force per square centimetre is H_0. The tension of one line of force is $S = H_0/8\pi$. It is well known

that the velocity of a wave running on a string is $V = \sqrt{(S/m)}$. If we substitute the values of S and m, we find $V = H_0/\sqrt{(8\pi\varrho)}$. The difference from (18.52) is explained by the fact that we have not taken into account the sideways pressure of the lines of force, which is also $H/8\pi$ (on one line). However, we shall not give the corresponding calculation here.

We can now estimate the retardation time of a mass of gas of diameter l, with density ϱ, moving in a field H_0 with velocity v. We can consider this cloud approximately as half a wave moving in one direction, like a rising crest. The wavelength is then $2l$, the velocity $v = H_0/\sqrt{(4\pi\varrho)}$, and consequently the period $T = 2l/v$. The retardation time of the cloud is

$$t \approx \tfrac{1}{4}\, T = l/2v = l\sqrt{(4\pi\varrho)}/2H_0 \,.$$

If $l = 1000$ km, $\varrho = 10^{-12}$ g/cm³, $H_0 = 10$ gauss, then $t \approx 20$ seconds. Thus, if we take self-induction into account, the retardation time is considerably increased.

If the conductivity is not infinitely large, the current j will lose energy by heating the matter, and the wave will be damped. If the damping is not too strong, the mean energy loss in 1 second is, by formula (18.37),

$$\begin{aligned}\frac{\mathrm{d}W}{\mathrm{d}t} &= -\frac{1}{T}\int_0^T \frac{1}{\lambda}\, j^2\,\mathrm{d}t = -\frac{A^2 c^2 \Omega^2 \varrho}{4\pi H_0^2 \lambda}\,\frac{1}{T}\int_0^T \cos^2\Omega\left(t-\frac{z}{v}\right)\mathrm{d}t \\ &= -\frac{A^2 c^2\Omega^2\varrho}{8\pi H_0^2\lambda}\,.\end{aligned} \qquad (18.56)$$

The wave energy W is composed of the magnetic energy $H_y^2/8\pi$ and the kinetic energy $\frac{1}{2}\varrho v^2$. By means of (18.53) we easily find that these two components are equal, so that the total mean energy is

$$W = \frac{1}{T}\int_0^T \frac{H_y^2}{4\pi}\,\mathrm{d}t = \frac{A^2}{4\pi}\,\frac{1}{T}\int_0^T \sin^2\Omega\left(t-\frac{z}{v}\right)\mathrm{d}t = \frac{A^2}{8\pi}\,. \qquad (18.57)$$

We introduce the quantity α defined by the relation

$$2\alpha = -\frac{1}{W}\frac{\mathrm{d}W}{\mathrm{d}z} = \frac{1}{W}\,\frac{1}{v}\,\frac{\mathrm{d}W}{\mathrm{d}t}\,.$$

Then $W = W_0\, e^{-2\alpha z}$. Since W is proportional to H_y^2,

$$H_y = A e^{-\alpha z}\sin\Omega(t - z/v)\,.$$

The logarithmic damping decrement is easily calculated from its definition and from (18.56) and (18.57) to be

$$\alpha = \tfrac{1}{2}\,\pi\,\frac{c^2}{\lambda}\,\frac{\Omega^2\,\varrho^{3/2}}{H_0^{\,3}}\,.$$

If the conductivity is anisotropic, then we must use λ_1 instead of λ, since only j_I evolves heat. The wave damping is important only for comparatively rapid oscillations whose periods are of the order of minutes.

7. The propagation of a magnetic field. Magneto-hydrodynamic waves form one of the mechanisms whereby a field is propagated in a conducting medium. If the medium cannot move, the field is propagated in the form of electromagnetic oscillations which are damped, but periodic, if the conductivity is small, or aperiodic if the conductivity is large.

In the latter case the field penetrates into the conducting medium to approximately $\frac{1}{6}$ of its wavelength in this medium. Consequently, a more slowly changing field penetrates deeper into the conductor. This produces what is called *electromagnetic screening* of the internal parts of the conductor from the external magnetic field; we have mentioned this above. The energy of the field is thereby transformed into the energy of the system of screening currents.

The problem of the propagation of a variable magnetic field is very important in the physics of the Sun, since the poles of this field (the spots) are localised at separate points on the Sun's surface, and the whole nature of the various phenomena depends on whether the field passes to the corona or to the prominences, and by how much its intensity diminishes. It is difficult to measure the field in a prominence or in the corona directly, because its intensity is small.

P. E. Kolpakov and Ya. P. Terletskiĭ [65] have indicated the important part played by the mobility of the solar atmosphere. We have seen previously that in this case the matter is "frozen" to the lines of force, there are no induction currents, and the energy of the field is not converted into heat; the field must consequently be propagated. This phenomenon may be more intelligibly interpreted as the motion of matter under the action of the magnetic pressure gradient together with the field, or as the propagation of the field in the form of magneto-hydrodynamic waves.

A second method of field propagation which is important for the Sun is the transmission of the field inside a large cloud. If there is a magnetic field inside the cloud, and the field outside disappears, then the surface layers of the cloud will cause electromagnetic screening of the outer layers, preventing the rapid propagation of the field and diminishing its intensity in the cloud.

Let us consider this phenomenon in greater detail, following T. G. COWLING. Since the field **H** has no sources, we can introduce a vector potential of the magnetic field, **A**, defined by the condition **H** = **curl A**. Substituting the value of **H** in (18.43) and changing the order of differentiation with respect to co-ordinates and time, we find

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} + \mathbf{grad}\,\phi\,, \tag{18.58}$$

where **E** is the electric field which appears as the magnetic field decays, and ϕ is a scalar potential. The latter term can be omitted, since no considerable volume charge can be formed in a conducting medium. Thus we find for the induction current density

$$\mathbf{j} = \lambda\,\mathbf{E} = -\frac{\lambda}{c}\frac{\partial \mathbf{A}}{\partial t}\,. \tag{18.59}$$

This current produces a magnetic field determined by equation (18.42). Replacing **H** on the left-hand side of (18.42) by **curl A**, and **j** on the right-hand side by (18.59), we obtain

$$\mathbf{curl\,curl\,A} = \mathbf{grad}\,\mathrm{div}\,\mathbf{A} - \nabla^2\mathbf{A} = \frac{4\pi\lambda}{c^2}\frac{\partial \mathbf{A}}{\partial t}\,. \tag{18.60}$$

Only **curl A** has hitherto been defined, and we can take any desired value for div **A**. We put div **A** = 0.

$$\nabla^2\mathbf{A} = \partial^2\mathbf{A}/\partial x^2 + \partial^2\mathbf{A}/\partial y^2 + \partial^2\mathbf{A}/\partial z^2$$

is not larger, in order of magnitude, than A/R^2, where R is the radius of the cloud (if $H \approx 0$ and $A \approx 0$ outside the cloud). Then (18.60) allows us to find the rate of decrease of **A**:

$$\partial\mathbf{A}/\partial t \leq -c^2\,\mathbf{A}/4\,\pi\lambda\,R, \tag{18.61}$$

and hence the decay time of **A** (and consequently of **H**) is

$$t \geq \frac{A}{-\partial A/\partial t} = \frac{4\,\pi\,\lambda\,R^2}{c^2}\,. \tag{18.62}$$

The time of growth of the field is of the same order of magnitude. If $R = R_{\odot} = 7\times 10^{10}$ cm, $\lambda = 10^{16}$ per second (the conditions in the interior regions of the Sun), then $t \approx 10^9$ years. This time is really determined by the self-induction of the volume concerned, and is the time to establish a steady state.

A. YA. KIPPER has pointed out that the presence of chaotic motions which distort the lines of force leads to a more rapid damping of the field. Such motions may arise under the action of magnetic pressure gradients formed when the original field decays inhomogeneously.

It should be emphasised that t characterises the variation of the flux through a material contour. The field intensity at a given point may vary rapidly, as we see from the example of magneto-hydrodynamic waves.

Thus, if the Sun acquired its magnetic field 10^9 years ago, it would have retained it up to the present time even in the absence of any "maintaining" effects. For a cloud with $R = 10^9$ cm and $\lambda = 10^{13}$/second (the conditions in the surface layers), $t \approx 300$ years. Thus the magnetic flux of a spot should exist without further energy loss for hundreds of years, while a spot in fact appears and disappears in the course of a few days. We shall consider a possible explanation of this in the next chapter.

Let us now consider in more detail the "freezing" of the field to the matter. Suppose that some material contour is drawn in a conducting medium. The variation in the magnetic flux embraced by the contour causes self-induction currents in it, which, according to Lenz's Law, oppose the change in the flux. If the conductivity is infinite, the magnetic flux through the contour does not change. In an actual conductor, the self-induction currents gradually decay; their energy is transformed into heat, and the flux gradually changes. The order of magnitude of the time during which the change takes place is given by (18.62), i. e. it is very large in cosmical conditions (in nebulae it is of the order of 10^9 years). During this time the contour can move with the matter and be deformed in any manner: it will carry the magnetic flux with it. Since the flux can be represented by the lines of force, the great time needed for the flux to vary means that the lines of force always remain within the limits of the material contour in question, i. e. they pass through the same body of material and move with the same velocity as they follow its motion, as for instance in the case of magneto-hydrodynamic waves.

If there are random motions in a conducting medium, the lines of force will have to take a very complicated form as they follow the matter, and will become "entangled". It is easy to see that the lines must thereby become more dense, i. e. the field intensity must increase. The total number of lines of force is unchanged, but any one line will intersect many times a plane through the gas in any direction. This increase in the field intensity cannot continue indefinitely. The field exerts a reciprocal action on the conducting medium (the magnetic pressure $H^2/8\pi$), and, when the mean density of magnetic energy is comparable with the mean kinetic energy density $\frac{1}{2}\varrho v^2$, the field ceases to increase and equilibrium is reached between the magnetic and kinetic energies. This mechanism of growth of the field is apparently of great importance in the interstellar medium, where we may expect a field of the order of 10^{-5} gauss to be present.

Chapter 19. Sunspots and faculae

1. Introductory remarks. Radiative equilibrium in sunspots. The following must be reckoned among the fundamental problems of the physics of sunspots:

(1) The study of the physical state of the matter inside spots, and the establishment of the nature of the thermal equilibrium and the depth distribution of the basic physical parameters (ionisation and excitation of atoms, etc.) in the spot.

(2) The elucidation of the factors which maintain a lower temperature in sunspots than in the surrounding layers of the photosphere.

(3) The origin of the strong magnetic fields in the spot and the study of the law of their variation with time.

(4) The general problem of the origin of spots and questions relating to the eleven-year cycle of solar activity. (The cause of the cycle itself, the elucidation of the heliographic regularities of the sunspot distribution at various phases of solar activity, the magnetic regularities connected with the phase of solar activity, such as the periodic interchange of polarity in the two hemispheres of the Sun, and so on.)

At present, fairly definite conclusions have been reached only as regards the first of these problems.

Let us consider the question of the physical state of the matter inside sunspots, and firstly the question of the temperature distribution within the spots. Spectroscopic observations have disclosed regular motions of the gases in spot regions. In the lower layers of the spot, the matter flows out of it, while in the upper layers the matter flows into it (the Evershed effect). The presence of such currents leads immediately to the idea that processes of heat transfer by moving matter, i. e. processes of **convective heat transfer,** might play an important part inside the spots. However, a number of facts indicate that, despite the presence of these currents, the heat transfer in a spot, as in the photosphere, takes place mainly by radiation. We shall briefly discuss these facts.

As in the case of the solar photosphere, the absorption processes in a spot whose spectral class is K 0 should be determined principally by negative hydrogen ions (see Chapter 6). But in this case, as was stated in Chapter 6, we can use the dependence of T on τ for grey material to determine the temperature distribution. The fact that the energy distribution curve in the continuous spectrum of the spot is approximately Planckian indicates that the continuous absorption coefficient in the spot varies little with frequency.

Consequently, if the photosphere and the spot are in a state of radiative equilibrium (we have already proved this for the photosphere), the total intensity distribution $I(\theta)$ of the emergent radiation as a function of the angle θ to the normal can be found with quite sufficient accuracy by the use of formula (4.32), both for the photosphere and for the spot. It then follows from this formula that the ratio $[I(\theta,0)]_s/[I(\theta,0)]_\odot$, where the suffix s refers to the spot, should be independent of the angle θ. In other words, the ratio of the total (bolometric) intensity of radiation for the spot to the total intensity for the neighbouring photosphere should be the same, for the case in question, at all points of the Sun's disc. Observations which have been made, including the investigations of G. F. Sitnik [150], entirely confirm the supposition that the ratio $[I(\theta,0)]_s/[I(\theta,0)]_\odot$ is constant.

The fact that this ratio is independent of the angle θ makes it possible to determine the effective temperature $(T_e)_s$ of a spot at any point of the solar disc. If $(T_e)_\odot$ is the temperature of the photosphere, then by (4.21)

$$(\pi H)_s/(\pi H)_\odot = [(T_e)_s/(T_e)_\odot]^4 . \tag{19.1}$$

Starting from observations of the ratio $[I(\theta,0)]_s/[I(\theta,0)]_\odot$, which is independent of θ, and is therefore equal to $(\pi H)_s/(\pi H)_\odot$, and knowing $(T_e)_\odot$, we obtain $(T_e)_s$. Such observations (taking scattered light into consideration) give a value of about 0·4 for the left-hand side of (19.1). With $(T_e)_\odot = 5710°$, this leads to $(T_e)_s$ about 4500°. The latter figure is a kind of average, since observation shows that the larger spots apparently have a smaller $(T_e)_s$. Thus, for example, measurements of an exceptionally large spot which was observed in November 1938 gave a temperature of about 3700°.

The values of $(T_e)_s$ found from formula (19.1) should not be identified, as is often done, with the excitation temperature $(T_{ex})^s$, since, as we have seen in Chapter 12, there is some difference, whose nature is not yet clear, between the two quantities.

A second method of solving the problem of heat transfer is the following. Assuming that the spot and the photosphere are in radiative equilibrium, and taking into account the fact that in both cases the absorption coefficient depends only slightly on the frequency, we can apply formula (4.40) to the spot and to the photosphere. In this case we have

$$\frac{[I_\lambda(\theta,0)]_s}{[I_\lambda(\theta,0)]_\odot} = \frac{\int\limits_0^\infty e^{-t\sec\theta}\sec\theta\,dt/[\exp\{hc/k\lambda(T_e)_s(\frac{1}{2}+\frac{3}{4}t)^{\frac{1}{4}}\}-1]}{\int\limits_0^\infty e^{-t\sec\theta}\sec\theta\,dt/[\exp\{hc/k\lambda(T_e)_\odot(\frac{1}{2}+\frac{3}{4}t)^{\frac{1}{4}}\}-1]} \tag{19.2}$$

If we apply it to the centre of the solar disc, where $\theta = 0$, the result is

$$\frac{[I_\lambda(0,0)]_s}{[I_\lambda(0,0)]_\odot} = \frac{\int_0^\infty e^{-t}\,\mathrm{d}t/[\exp\{h\,c/k\,\lambda\,(T_e)_s\,(\frac{1}{2}+\frac{3}{4}\,t)^{\frac{1}{4}}\}-1]}{\int_0^\infty e^{-t}\,\mathrm{d}t/[\exp\{h\,c/k\,\lambda\,(T_e)_\odot\,(\frac{1}{2}+\frac{3}{4}\,t)^{\frac{1}{4}}\}-1]} \tag{19.3}$$

Thus formula (19.3) gives the theoretical ratio of intensities "spot to adjoining photosphere" for various wavelengths, for a spot at the centre of the Sun's disc. If the spot is in radiative equilibrium, formula (19.3) can be brought into agreement with the observed ratio $[I_\lambda(0,0)]_s/[I_\lambda(0,0)]_\odot$ by choosing the one parameter $(T_e)_s$, if, of course, the value of $(T_e)_\odot$ is given. This can in fact be done for $(T_e)_s \approx 4500°$, the observed intensity ratio then being in satisfactory agreement with the theoretical ratio in the wavelength range from 3000 Å to 10,000 Å; the agreement is not quite so good from 10,000 Å to 22,000 Å, possibly because of the effect of water-vapour absorption bands in the Earth's atmosphere. Further observations are needed here.

Finally, we can apply formula (19.2), for any wavelength λ, to investigate the ratio $[I_\lambda(\theta,0)]_s/[I_\lambda(\theta,0)]_\odot$ at various points on the solar disc (as in the case of integrated radiation). Here too the theory is in satisfactory agreement with observation; the values of $(T_e)_s$ obtained in this way vary with different observers, from 4300° (R. S. Richardson) to 4750° (G. F. Sitnik). It is possible that this difference is related to an actual difference in $(T_e)_s$; as we have already said, larger spots apparently have smaller $(T_e)_s$.

From these facts we may suppose that sunspots are, in fact, in radiative equilibrium. In particular, the assumption that the spots are in convective equilibrium leads to a marked discrepancy between theory and observation.

The fact that the matter in a spot must be in radiative equilibrium follows also from an estimate of the energy fluxes carried by radiation and by convective currents. At the velocities observed in a spot (up to 2 or 3 km/sec), convective currents are too small to be an effective agent for thermal energy transfer*.

2. The spectrum of a spot. The temperature. The electron pressure. We shall now give an analysis of the absorption lines in the spectrum of a sunspot. This spectrum, in accordance with the lower temperature of the spot, belongs to a later class than the spectrum of the photosphere; spots are generally placed in the class K 0, while the Sun's spectrum is in the class dG 3. If we suppose that the physical conditions in a spot

* A detailed survey of all the topics discussed above is given by G. F. Sitnik [151].

are similar to those existing in the atmosphere of a star of the class K 0, and take into account the fact that the acceleration due to gravity in the spot must be the same as for the Sun (a dwarf star) then we should expect, from Table 11, that $(T_e)_s$ would be 4910°. This is somewhat higher than the figure given above (4500°). However, no reliable conclusions can yet be drawn, since the class K 0 for a spot is approximate only. Moreover, we should, strictly speaking, use for comparison the *mean* spectrum of the spot, i. e. the spectrum which a distant observer would record if the entire disc of the Sun were covered with spots. If, for example, the class K 0 mentioned corresponds to a spot at the centre of the Sun's disc, the *mean* spot spectrum will be of a later* class, and the agreement with $(T_e)_s \approx 4500°$ will be improved.

In studying the excitation and ionisation conditions in spots, we can use all the methods explained in Part II. That is, we can use either the elementary methods of the curve of growth (Chapter 12) or those of Chapter 13, which depend on a more accurate construction of contours. The construction of a curve of growth for sunspots is carried out in accordance with the general methods explained in Chapter 12. The existing investigations show that the curve of growth thus obtained differs hardly at all in shape (within the limits of observational error) from the curve of growth for the undisturbed photosphere. However, the data required for more definite conclusions are as yet scanty. Any displacement and distortion of the curve of growth for spots, in comparison with that for the photosphere, should be mainly due to three factors: (1) the general difference in the number of absorbing atoms in the spot and in the photosphere; (2) the difference in temperature between spot and photosphere; (3) the splitting of the lines in the spot spectrum owing to the presence of a strong magnetic field (up to 4000 gauss) in the spot.

If the resolving power of the spectrograph is not very high, the separate Zeeman components merge together, and this is equivalent to a broadening of the lines. The effect of such a broadening on the equivalent widths of the lines should be unimportant for faint lines and most considerable for *medium* lines, corresponding to the flat transitional part of the curve of growth.

Let us consider these cases in turn. For faint lines, the equivalent width is proportional to the number of absorbing atoms, and since this number remains the same in the presence of a field, the effect of the splitting amounts merely to some broadening of the lines, without any change in W_λ. For very strong lines, the magnetic splitting, even for a field of the order of 4000 gauss, is considerably less than the extent

* Since the radiation of the central parts of the disc of the Sun (or of a star) corresponds to hotter layers than that from the disc as a whole.

of the wings, and usually less than the width of the line (say for $r_\nu = 0{\cdot}5$), even taking into account the outermost components.

We now consider medium lines. Let the equivalent width of some line lying on the intermediate part of the curve of growth be $W_\lambda{}^0$ in the absence of a magnetic field, and let the number of atoms producing it be N. In the presence of a magnetic field the line is split into n components, each component being determined, roughly speaking, by the absorbing action of $N_s = N/n$ atoms. However, for the lines considered (on the transitional part of the curve of growth), even a large change in the number of absorbing atoms causes a very small change in the equivalent width. Hence, for small n, each of these components should have practically the same W_λ as the original line had in the absence of the magnetic field, i. e. $W_\lambda{}^0$. The merging of all the components results in an increase of W_λ for the observed line.

The polarisation of the atoms leads to a certain decrease* in this resultant width, but the line will nevertheless be strengthened if the number n of components and the field strength are sufficiently great. Thus the Zeeman effect leads to a raising of the middle part of the curve of growth. Here it must be borne in mind that different atoms (and even different lines belonging to the same atom) give different splitting schemes, both as regards the magnitude of the splitting and in the relative intensities of the components. Hence the problem in question is extremely complicated.

Curves of growth constructed both for the spot and for the photosphere allow a number of conclusions to be drawn concerning the physical conditions in the spot. Firstly, we can determine the excitation temperature in the spot. We first determine a *provisional* value of $(T_{\text{ex}})_s$ (in constructing the curve of growth), and then, using formula (12.37) and a graph like Fig. 34, we find an *improved* value of $(T_{\text{ex}})_s$.

Another method is based on formula (12.40), in which we insert the suffix s instead of st:

$$(\log_{10} X_0)_s - (\log_{10} X_0)_\odot = \log_{10} \frac{(N_r)_s}{(N_r)_\odot} \frac{(u_r)_\odot}{(u_r)_s} - \log_{10} \frac{(v_0)_s}{(v_0)_\odot} - 5040\, \varepsilon_i \left[\frac{1}{(T_{\text{ex}})_s} - \frac{1}{(T_{\text{ex}})_\odot} \right]. \qquad (19.4)$$

By determining the difference $\{(\log_{10} X_0)_s - (\log_{10} X_0)_\odot\}$ for various multiplets whose lower levels have different ε_i, we can construct the

* The effect of the polarisation of the atoms is that the energy absorbed in each of the separated components is less than it would be if there were no polarisation.

relation between these differences and the values of ε_i, which is usually linear, in agreement with (19.4). The quantity

$$\Delta\theta = 5040\left[\frac{1}{(T_{\text{ex}})_s} - \frac{1}{(T_{\text{ex}})_{\odot}}\right] \qquad (19.5)$$

is then found from the angle between the straight line obtained and the axis of ε_i. Having found $\Delta\theta$ and taken a definite value for $(T_{\text{ex}})_{\odot}$, we obtain $(T_{\text{ex}})_s$. For large spots, P. TEN BRUGGENCATE and H. VON KLÜBER have found $\Delta\theta = 0{\cdot}33$. Taking this value with $(T_{\text{ex}})_{\odot} = 5040°$ we obtain $(T_{\text{ex}})_s = 3800°$, while with $(T_{\text{ex}})_{\odot} = 5700°$ we obtain $(T_{\text{ex}})_s = 4200°$.

Apparently $(T_{\text{ex}})_s$, like $(T_e)_s$, varies from one spot to another, being less for larger spots. Hence the value mentioned, $\Delta\theta = 0{\cdot}33$, is probably too high.

The methods based on the theory of curves of growth also allow us to determine the electron pressure in a spot. Having found the number of atoms of some element in two successive ionisation states, we must use the ionisation formula (5.11), taking, of course, a definite value for T which is close to the effective temperature of the spot. If we do this for two or more elements, then we can determine not only p_e, but also T itself.

The electron pressure p_e in a spot can also be obtained without using curves of growth. Writing the ionisation formula (5.19) twice (for the photosphere and for the spot), we find, putting the result in logarithmic form,

$$\log_{10}\left(\frac{N_{1,1}}{N_{0,1}}\right)_s = \tfrac{5}{2}\log_{10}T_s - \frac{5040}{T_s}\chi_0 - \\ - \log_{10}(p_e)_s + \log_{10}\frac{g_{1,1}}{g_{0,1}}\frac{2(2\pi m_e)^{3/2}k^{5/2}}{h^3}, \qquad (19.6)$$

$$\log_{10}\left(\frac{N_{1,1}}{N_{0,1}}\right)_{\odot} = \tfrac{5}{2}\log_{10}T_{\odot} - \frac{5040}{T_{\odot}}\chi_0 - \\ - \log_{10}(p_e)_{\odot} + \log_{10}\frac{g_{1,1}}{g_{0,1}}\frac{2(2\pi m_e)^{3/2}k^{5/2}}{h^3}, \qquad (19.7)$$

where, of course, we assume that equation (5.19) can be applied to the atmosphere as a whole (passing from n to N). Subtracting one equation from the other, we obtain

$$\log_{10}\frac{(p_e)_s}{(p_e)_{\odot}} = \log_{10}\frac{(N_{1,1})_{\odot}}{(N_{1,1})_s} + \log_{10}\frac{(N_{0,1})_s}{(N_{0,1})_{\odot}} + \\ + \frac{5}{2}\log_{10}\frac{T_s}{T_{\odot}} - 5040\,\chi_0\left(\frac{1}{T_s} - \frac{1}{T_{\odot}}\right), \qquad (19.8)$$

where the temperatures which appear here are, strictly speaking, some ionisation temperatures averaged with respect to depth, which can be approximately identified with the effective temperatures.

The first two terms on the right-hand side of (19.8) can be determined if two successive ionisation states of some element are represented in the spectra of the spot and of the photosphere by fairly strong resonance lines, for which, according to formula (12.28), W_λ is proportional to $\sqrt{N}$, and for which the use of the curve of growth is consequently not obligatory*.

The determination of $(p_e)_s$ from formula (19.8), using the resonance line 4227 Å of Ca I and the H and K lines of Ca II, has been carried out by T. V. KRAT [68]. The value 3800°, mentioned above, was taken for T_s, and 5040° for $T_\odot$. The application of the law $W_\lambda \sim \sqrt{N}$ as above gave $(p_e)_s/(p_e)_\odot \approx 1/40$ from measured equivalent widths. This estimate of $(p_e)_s$ is certainly too low, since the difference $\Delta\theta$ taken by T. V. KRAT is above the average. However, on using other existing values for this difference, we invariably arrive at the conclusion that $(p_e)_s$ is about an order of magnitude less than $(p_e)_\odot$.

Let us compare this difference in p_e between the photosphere and the spot with that which we should expect in passing from the Sun to a star of the class dK 0. It follows from Table 8 that, in the atmosphere of the star 70 Ophiuchi A, of class dK 0, the electron pressure is about 20 times less than in the atmosphere of the Sun. This difference agrees in order of magnitude with T. V. KRAT's estimates.

Furthermore, a direct comparison of this star 70 Ophiuchi A with a sunspot, carried out by S. E. A. VAN DIJKE [170], has shown that, within the limits of observational error, the values of p_e are practically the same for the two objects. These facts indicate that a spot is similar in many respects to the atmosphere of a dwarf star of the same spectral class.

It would be interesting to confirm this last conclusion on the basis of other considerations. It would in principle be possible to conduct a quantitative comparison of equivalent widths in the spectra of the spot and of the photosphere, using the curve of growth in order to convert to numbers of atoms. However, a number of complications immediately occur here. Thus, for instance, the number of excited magnesium atoms which produce the well-known green triplet 5183·7 Å, 5172·7 Å and 5167·4 Å in the spot is twice the number in the photosphere. Calculations show that this result is very difficult to explain

* It must be remarked that the use of this law presupposes that the damping constant Γ_{ik} is the same. This constant Γ_{ik} is determined chiefly by collision damping, and consequently may be different in the spot and in the photosphere. The method in question is therefore only approximate.

by using only the ordinary formulae of thermal equilibrium. Moreover, the increase in the equivalent widths of these lines as we pass from the photosphere to the spot is practically the same as in passing from dG 3-type stars to dK 0-type stars.

T. V. KRAT thinks that the cause of these discrepancies is to be sought in the fact that the continuous absorption coefficient in the spot should be less than in the photosphere. This increases the value of $\eta_\nu = \sigma_\nu/\varkappa_\nu$ in the spot, and thereby has the same effect as an increase in the number of atoms. The fact that $\varkappa_\nu$ decreases from photosphere to spot has been established by T. V. KRAT, by a study of the deviations in the intensity of radiation in the continuous spectrum of spot and photosphere from the intensity of radiation corresponding to the case of grey material [66]. The inequality $(\varkappa_\nu)_s/(\varkappa_\nu)_\odot < 1$ also follows from Fig. 15, if we take $T_s = 4200°$, $T_\odot = 5600°$ and calculate the ratio $(\varkappa_\nu)_s/(\varkappa_\nu)_\odot$ for the value of $(p_e)_\odot/(p_e)_s = 10$ given above. For $\lambda = 5000$ Å, the ratio $(\varkappa_\nu)_s/(\varkappa_\nu)_\odot$ is found to be about one-third.

It follows from these considerations that, to give a theoretical interpretation of the differences in line intensities between spot and photosphere, we must henceforward use the methods of Chapter 13. Having found the distribution of the various physical parameters inside the spot, we must construct the contours of the lines in which we are interested, in accordance with the general ideas of Chapter 13. The same, of course, must be done for the photosphere. Only then will the comparison of theory and observation be sufficiently justified.

For the calculations just mentioned concerning the structure of the spot, we can use the theory of the solar photosphere discussed at the beginning of Chapter 17. Here we need only take $T_e \approx 4500°$, retaining the same value of $g = 2{\cdot}74 \times 10^4$ cm/sec^2. It is possible, however, to proceed in a different manner, using the observationally determined law of variation of the intensity with the angle θ for the spot. This gives (see Chapter 17) the relation between B_λ and τ_λ. Using this method, R. MICHARD [87] finds that, in the range from 4800 to 6000 Å, the absorption of radiation in the spot is due to negative hydrogen ions. The model which he has constructed makes it possible to calculate the energy distribution in the infra-red continuous spectrum of a typical spot, and the calculations (based on the H^- ions) are in very good agreement with the results of E. PETTIT and S. B. NICHOLSON. The model accounts equally well for the line spectrum of the spot.

If we know how the temperature varies, and also the electron and gas pressures as functions of optical depth, we can calculate the geometrical depth of the spot and of the photosphere. A comparison shows that at equal depths the pressure in the spot is much less than in the photosphere (see above).

In such a comparison, the choice of the zero of linear depths in the spot and in the photosphere is very uncertain. This problem has recently been discussed by V. S. Berdichevskaya [21]. She finds that, for a particular adjustment of this common zero-point, the calculated values of the gas pressure in the spot may even come out greater than at the corresponding levels in the photosphere.

A second extremely important method of investigating the physical conditions in sunspots is a careful comparison of the spectrum of a spot with those of stars of neighbouring classes. We must find whether it is possible to choose, among the spectra of stars (of course, near the class dK 0), one which *exactly* corresponds to the spot spectrum. The only comparison of this kind is that of a spot spectrum with the spectrum of the dK 0-type dwarf 70 Ophiuchi A (see above), and this shows that the ratio of the number of atoms in the spot to the number in this star is greater than unity for all neutral atoms and less than unity for all ionised atoms.

3. Possible causes of the lower temperature in a spot. The magnetic fields in sunspots. The next question in the physics of sunspots is: what is the nature of the mechanism which decreases the temperature of the matter in a spot? Numerous attempts to explain this fact have been based on the assumption that an outflow and adiabatic expansion of gases takes place in a spot, and this leads to cooling. However, it has subsequently been discovered that this "adiabatic theory" meets with a number of very serious difficulties [179, p. 194]. More recent theories relate the cooling of the matter in spots to the presence of strong magnetic fields in them. We shall now proceed to consider the subject of the magnetic fields in spots.

We first give some general characteristics of the magnetic fields in sunspots. The presence of a strong magnetic field is an invariable property of sunspots. No spot has been observed without a field. On the other hand, we even find what are called **invisible spots**, which are regions of the photosphere above which weak magnetic fields are observed. These regions are often found where a spot has recently been or where one soon appears.

The magnetic lines of force are perpendicular to the Sun's surface at the centre of the spot and make an angle of about 20° with it at the edge of the penumbra. A single spot resembles a unipolar magnet in the nature of its field, while a typical group of two spots resembles a bipolar magnet. Sometimes multipolar groups are found, with a complex disposition of the poles.

A *bipolar group* can be characterised by its magnetic moment $\mathrm{M} = \mathrm{H} \cdot S\,l$, where H is the mean field intensity in the spot, S its

area and l the distance between the spots. The field intensity of a bipolar group at a distance R from it such that $R \gg l$ is

$$\mathbf{H} = -\operatorname{grad}\frac{\mathbf{M}.\mathbf{R}}{R^3} = \frac{3(\mathbf{M}.\mathbf{R})\mathbf{R}}{R^5} - \frac{\mathbf{M}}{R^3}, \tag{19.9}$$

i. e. the field decreases approximately as R^{-3}. The field intensity H at the centre of a spot varies with the area S of the spot, reaching 3500 gauss for large spots.

The dependence of H and S on time is shown in Fig. 50 for typical spots which change rapidly (below) and slowly (above) [35]. S is given in millionths of a hemisphere of the Sun, and H in gauss. At first H is proportional to S, and then reaches a maximum where it scarcely changes. Only later, as the area of the spot appreciably decreases, does H begin to decrease also.

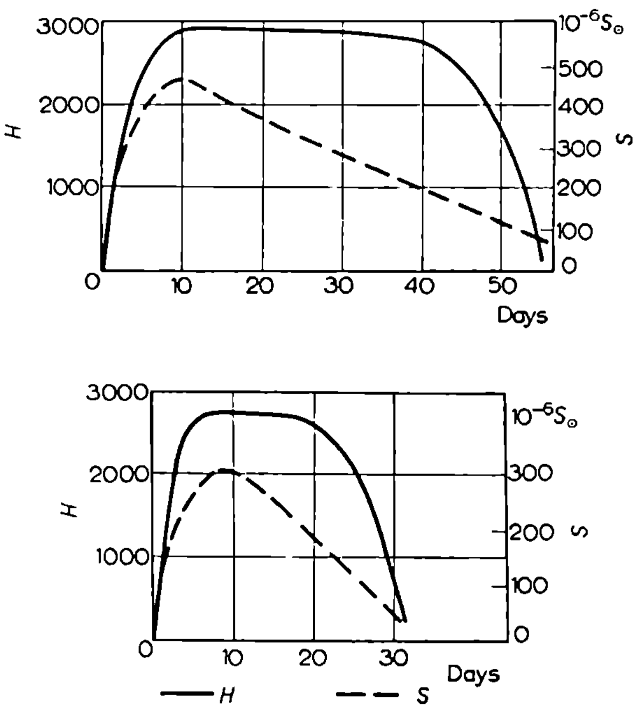

Fig. 50

The time of increase of the field intensity is usually 2 or 3 days; the time of constant field intensity lasts from some days to some weeks, and the time of decrease is one or two weeks. On the other hand, we have seen at the end of Chapter 18 that the field of a spot can exist for hundreds of years without loss of energy. Because of this, T. G. Cowling supposes that the spot field is carried out to the surface and down again by the motion of ionised matter. The spots are like windows, through which the field which always exists in the sub-photospheric layers emerges to the surface.

The magnetic field exerts a pressure $H^2/8\pi$ on the conducting atmosphere. In the deep layers, the gas pressure is considerably greater than the magnetic pressure, and hence the magnetic pressure does not lead to an expansion of the gas. In the *photospheric* layers, the magnetic pressure is of the same order of magnitude as the gas pressure, and the field may reach the surface by causing the gas to move. The expansion of the gas results in its cooling. In the equilibrium state the total pressure in the spot (gas pressure plus magnetic pressure) must be equal to the gas pressure outside the spot at the same level. Thus,

knowing H, we can estimate the difference between the gas pressure in the spot and that in the photosphere. Since $p = n\,kT$, the knowledge of the temperature ratio obtained from observation enables us to find the density ratio.

The radiation from the surrounding layers continually heats the outermost layers of the spot, increasing the gas pressure. The equilibrium is thereby destroyed and the heated layers expand somewhat, until their temperature falls to its former value. It is possible that this expansion is in fact the observed motion of matter from the spot.

Thus the lower temperature in a spot is apparently a necessary consequence of the strong magnetic field. There is as yet no conclusive answer to the problem of the origin of the field. L. É. Gurevich and A. I. Lebedinskiĭ [53] suppose, in opposition to Cowling, that the field can be directly produced in the sub-photospheric layers by the circulation of conducting matter in a magnetic field under the action of hydrodynamical forces. Such a circulation could cause currents which would determine the magnetic field of the spot. The whole process is reminiscent of the way in which dynamos are excited. However, the authors give no specific mechanism for this self-excitation.

Recent investigations with the solar magnetograph (see H. W. Babcock [13]) seem to show with certainty that the Sun has a weak general magnetic field. At latitudes below $\pm 65°$, regions have also been discovered which show the presence of weak magnetic fields. However, these regions are very impermanent and are transitory in character.

An important question is whether or not the origin of the magnetic field H in spots is related to the existence of the general magnetic field of the Sun. This question has been considered by V. E. Stepanov [163]. If the appearance of spots is connected in some way with the general field of the Sun, the place where the spots appear should depend mainly on the *magnetic*, and not the *heliographic*, latitude. Thus, if the Sun's magnetic pole in the northern hemisphere is, at a given moment, closer to us than the heliographic pole (the two do not coincide), the mean latitude at which spots appear in the northern hemisphere should be less than in the southern hemisphere. For the opposite position of the pole, the effect will be opposite.

The difference in the mean latitude at which spots originate in the two hemispheres should thus have a period equal to the period of revolution of the Sun's magnetic axis, as seen from the Earth. A statistical reduction of forty-five years' observations of spots has enabled this period to be fairly reliably determined: it is found to be 27·5993 days, which is somewhat different from that determined spectroscopically. However, the value obtained by V. E. Stepanov is more reliable,

since it is based on a lengthy series of observations. Thus it can be assumed that the appearance of spots is connected with the general field of the Sun.

4. Solar faculae. The depth distribution of temperature in a facula. Let us now consider the faculae. These, as is well known, are observed only on the outer parts of the solar disc, and chiefly in the region where $0{\cdot}6\, R_\odot < R < R_\odot$. They appear to the observer as brighter parts of the photosphere. Faculae have a complex fibrous structure, which sometimes changes in the course of a few hours. The fibres (filaments) are in turn composed of granules. However, the granulation in faculae differs from that in the undisturbed parts of the photosphere. The *granule-interstice* contrast in the faculae is considerably greater than in the photosphere.

In particular, the contrast at the extreme limb is, according to V. A. KRAT [73], about 30 %. Unlike ordinary photospheric granules, facular granules have a tendency to be formed at almost exactly the same places. This, apparently, explains the surprising stability of faculae, regarded as series of facular granules. According to recent investigations by V. A. KRAT and by C. MACRIS, it seems that the mean lifetime of a single granule in a facula is approximately the same as in the photosphere.

It seems that the mean level at which faculae lie is higher than that of spots and of the photosphere as a whole. This is suggested, for example, by the fact that the angular velocity of rotation of the Sun, as determined from faculae, is greater than that determined from spots. It is known that the angular velocity of rotation of the various layers of the Sun increases with height. (This apparently does not extend to the solar corona; see Chapter 22.)

The measurement of the contrast between the *facula* and the *adjoining photosphere* is a very important subject in the physics of faculae. The ratio

$$F_\lambda(\theta) = [I_\lambda(\theta, 0)]_f / [I_\lambda(\theta, 0)]_\odot \,, \tag{19.10}$$

where the suffix f refers to a facula, has been measured for various θ by many observers. Thus, V. A. AMBARTSUMYAN and N. A. KOZYREV found for a selected facula with $R = 0{\cdot}94\, R_\odot$ a value of $F(\theta)$ equal to 1·14 for $\lambda \approx 3900$ Å and 1·10 for $\lambda \approx 5000$ Å. The values of $F(\theta)$ have been determined by T. V. KRAT [67] from several faculae for various θ and for $\lambda = 3900$ Å and $\lambda = 5000$ Å. For $\lambda = 3900$ Å, the value of $F(\theta)$ for $R = 0{\cdot}52\, R_\odot$ and $R = 0{\cdot}90\, R_\odot$ was 1·15 and 1·28 respectively. For $\lambda = 5000$ Å the value of $F(\theta)$ for the same R was 1·08 and 1·16. A. I. LEBEDINSKIĬ and S. S. ZHURAVLEV [79] have studied the value of $F(\theta)$ by photoelectric methods.

It must be pointed out that all these investigations relate essentially, not to the facular filaments mentioned, but to some mean "smeared" representation of the facula. This smearing occurs because of the oscillation of the images and the presence of scattered light. On account of the oscillation of the images, the brightness of the filaments will be partly "distributed" on to the parts near the filaments. Calculations which take account of this effect, and also direct (provisional) estimates, suggest that the value of $F(\theta)$ in these filaments may in reality reach 1·5 to 2·5.

It is very interesting that the ratio of *total* intensities $[I(\theta,0)]_f/[I(\theta,0)]_\odot$ varies (unlike the ratio for sunspots) with the distance from the centre of the disc. This apparently means that *solar faculae are not in a state of radiative equilibrium.*

The temperature distribution inside a facula can be studied by means of the method explained at the end of Chapter 7. To do so, we must know the law of variation of $[I_\lambda(\theta,0)]_f$ with the angle θ. In particular, we can find this law from the observed ratio (19.10), since the dependence of $[I_\lambda(\theta,0)]_\odot$ on θ is known. Knowing the law of variation with θ of the intensity concerned, we can regard the equation

$$[I_\lambda(\theta,0)]_f = \int_0^\infty B_\lambda(T_f)\, e^{-\tau_\lambda \sec\theta} \sec\theta \, d\tau_\lambda \qquad (19.11)$$

as an integral equation which enables us to determine the dependence of $B_\lambda(T_f)$ on τ_λ for a given λ. If we suppose that the matter in the facula and in the photosphere has the properties of "grey material"*, we can apply the same method to the total radiation also:

$$[I(\theta,0)]_f \approx \int_0^\infty B(T_f)\, e^{-\tau \sec\theta} \sec\theta \, d\tau \,. \qquad (19.12)$$

The solution of equations (19.11) and (19.12) can be obtained by the method of numerical integration. If it is possible to represent the observed intensity in the form of a polynomial, expanded in powers of $\cos\theta$, we can use the method of Chapter 7.

Such calculations show that for small τ the temperature T_f of the facula is *higher* than that of the photosphere, while for large depths it is *lower* than that of the photosphere. [It is therefore incorrect to equate (as is frequently done) the difference in the temperature of the radiation T_E between facula and photosphere to the difference in their effective temperatures.]

* In view of the small difference between the mean temperatures of facula and photosphere.

In order to understand the result just stated, it is necessary to recall that at the limb of the solar disc we observe, in general, radiation emerging from the uppermost layers of the solar atmosphere. Hence the greater brightness of the faculae at the limb, compared with the brightness of the photosphere, is a direct indication that in their outer layers they are hotter than the photosphere. On the other hand, we do not see the faculae at the centre of the disc. This means that the excess of thermal radiation, just mentioned, in the outer layers of the faculae (in comparison with the photosphere) must be in some degree compensated by a lower temperature of the faculae in regions where τ is large.

Unfortunately, the results of the above-mentioned calculations are as yet purely qualitative. We have said in Chapter 7 that the finding of the function $T(\tau)$ from the darkening law is affected by uncertainties arising from the necessary extrapolation of the intensity to the extreme limb of the solar disc. This problem, moreover, is especially complicated for the faculae. It was supposed until quite recently that the facula-photosphere contrast increases right up to the limb. The investigations of M. Waldmeier [178], however, indicate that this contrast is greatest for $R_m \approx 0{\cdot}95\ R_{\odot}$ and that it decreases on each side of R_m. In the two cases a different distribution $T(\tau)$ is found. In the former case (maximum contrast at the limb of the disc) the excess heating in the facula should take place mainly in its outermost layers, with $\tau < 0{\cdot}02$. On the other hand, in the latter case the excess heating (i. e. the ratio $T_f/T_{\odot}$) is a maximum at about $\tau \approx 0{\cdot}6$.

To solve this problem, we need further careful measurements of the facula-photosphere contrast close to the extreme limb of the Sun. There is still reason to think that the second case seems closer to reality, since, as is shown by a study of absorption lines in the spectra of faculae, the atoms which form these lines are at a higher temperature than the corresponding atoms in the photosphere (see below), while the absorption lines are formed *on the average* close to the level $\tau = 0{\cdot}3$ to $0{\cdot}6$. The study of *flocculi* also indicates that the second case is closer to reality. We shall discuss this subject in Chapter 21.

In connection with the interpretation of the observed facula-photosphere contrast, we shall pause to consider an important problem which is quite general in character. We assume for simplicity that the radiation of frequency ν emerging from the Sun (or a star) arises in a relatively thin opaque surface layer where all the temperatures which characterise the radiation and the state of the matter in the layer are *the same*, do not vary with depth, and are equal to T_0. In this case it is quite evident that the intensity of the emergent radiation $I_\nu(0)$ in the given frequency will be equal to $B_\nu(T_0)$. [See, for example,

formula (3.38) for $T = T_0 =$ constant.] We now assume that, for some reason, the *kinetic* temperature of the layers concerned is raised to some value $T_\varepsilon > T_0$. The question is whether the intensity $I_\nu(0)$ will now be equal to $B_\nu(T_\varepsilon)$. We shall show that this will not be the case*.

In accordance with formulae (3.30), (3.13) and (2.27), the most general expression for the intensity of the emergent radiation at the surface has the form

$$I_\nu(\theta.0) = \int\limits_0^\infty \frac{j_\nu}{\varkappa_\nu} e^{-t_\nu \sec\theta} \sec\theta \, dt_\nu \,, \tag{19.13}$$

where we have not used the assumption of local thermodynamic equilibrium. Restricting ourselves to a qualitative treatment of the problem, we shall assume that within the outer layers of the atmosphere, from which the intensity $I_\nu(\theta,0)$ mainly originates, the ratio $j_\nu/\varkappa_\nu$ is constant. On the hypothesis of local thermodynamic equilibrium, this assumption would imply, by (3.10), the constancy of the temperature inside these layers. Putting $j_\nu/\varkappa_\nu =$ constant, we have for all θ, by (19.13),

$$I_\nu(\theta,0) = j_\nu/\varkappa_\nu \,. \tag{19.14}$$

Next, taking into account the fact that the emission coefficient j_ν in stellar atmospheres is principally determined by recombinations (except in the far infra-red region of the spectrum), we can use formula (8.38) for it, neglecting stimulated emission processes for the case of the Sun. Finally, let the absorption processes (the coefficient $\varkappa_\nu$) and the emission processes (the coefficient j_ν) be determined, *in the frequency considered, by the same atoms*. Then, according to (5.20) and (5.22), we can write

$$\varrho\, \varkappa_\nu = n_{r,k}\, k_\nu' \,, \tag{19.15}$$

where $n_{r,k}$ is the number of atoms per cm^3 in the rth ionisation state and the kth excitation state. Here it must be emphasised that only one of the bands in Fig. 11 is taken into account, both in (8.38) and in (19.15). Nothing essentially new would be added to our results by taking account of all bands with $\chi_{r,k} < h\nu$.

From the above, we find for $I_\nu(\theta,0)$ the expression

$$I_\nu(\theta,0) = \left(\frac{n_e\, n_{r+1}}{n_{r,k}}\right) \frac{g_{r,k}\, h^4 \nu^3}{u_{r+1}\, c^2 (2\pi m_e k T_\varepsilon)^{3/2}} e^{-(h\nu - \chi_{r,k})/kT_\varepsilon} \,. \tag{19.16}$$

If in this formula we introduce the formulae of thermal equilibrium given in Chapter 5, putting $T = T_\varepsilon$ and taking account of stimulated emission, then we find $I_\nu(\theta,0) = B_\nu(T)$, a result which is easily intelligible.

* This question is discussed by É. R. Mustel' [100].

In order to make our principal result clearer, let us consider a frequency ν close to or equal to the frequency ν_k of the edge of the absorption band which we are discussing. In this case $h\nu \approx \chi_{r,k}$, and consequently

$$I_\nu(0,0) \approx \left(\frac{n_e n_{r+1}}{n_{r,k}}\right) \left\{\frac{g_{r,k} h^4 \nu^3}{u_{r+1} c^2 (2\pi m_e k T_\varepsilon)^{3/2}}\right\}. \tag{19.17}$$

The first factor on the right-hand side of (19.17) depends on the ionisation and excitation processes, while the second depends only on the kinetic temperature, and only very slightly on that. If the ionisation and excitation processes are determined by the temperature $T = T_\varepsilon$, then, as we have just said, $I_\nu(0,0) = B_\nu(T)$.

Let us now consider what will happen when T_ε is raised. We have seen in Chapter 8 that, under the conditions existing in stellar atmospheres (low densities of matter), the ionisation is determined mainly by radiation, and not by collisions. Moreover, we have said in Chapter 15 that the same is true of excitation processes. It follows from this that the first factor on the right-hand side of (19.17) will not, in general, change at first when T_ε increases. The effect of the increase in T_ε on this factor becomes noticeable only when the factor α in formula (8.24) becomes comparable with or greater than unity. The effect of the increase in T_ε on the second factor in formula (19.17), on the other hand, is generally small. Thus, when T_ε increases, the value of $I_\nu(0,0)$ will at first remain practically constant. Later (when $\alpha > 1$) it begins to increase, but the inequality $I_\nu(0,0) < B_\nu(T_\varepsilon)$, or even $I_\nu(0,0) \ll B_\nu(T_\varepsilon)$, is still maintained. Thus, under the conditions existing in stellar atmospheres, an increase in the intensity of the emergent radiation cannot be immediately related to any corresponding increase in the kinetic temperature. Hence, if for example the temperature of the radiation of a facula $(T_E)_f$ is greater than the temperature of the radiation $(T_E)_\odot$ of the neighbouring parts of the photosphere, the difference $\Delta T_E = (T_E)_f - (T_E)_\odot$ cannot be ascribed to a corresponding difference in the kinetic temperatures. In reality, the difference $\Delta T_\varepsilon = (T_\varepsilon)_f - (T_\varepsilon)_\odot$ may be much higher. It is clear, also, that this conclusion is entirely qualitative and is not connected with the restrictions made in our discussion of the problem.

5. The spectra of faculae. We now turn to the spectra of faculae. A general qualitative comparison of the spectrum of a facula with that of the adjoining photosphere shows that the lines of ionised metals are stronger in the former than in the latter. The reverse is observed for the lines of neutral metals. The H and K lines of Ca II are an exception to this rule; they are fainter in the spectrum of a facula than in that of the adjoining photosphere.

There are as yet few spectrophotometric investigations of faculae. They are chiefly the work of T. V. KRAT [67] at Pulkovo and of O. N. MITROPOL'SKAYA [89] at the Crimean Astrophysical Observatory. We shall give briefly the results of these two investigations. The construction of the curve of growth for a facula, which was performed by O. N. MITROPOL'SKAYA using lines of Fe I, has shown that this curve is in shape the same, within the limits of observational error, as the curve constructed for the adjoining regions of the photosphere; a certain relative displacement of the two curves along the axis of abscissae shows that there are fewer absorbing atoms of Fe I in the facula than in the adjoining photosphere.

It follows from these two investigations (in agreement with the qualitative comparison mentioned above) that the great majority of the lines of neutral elements are in fact fainter in the facula than in the photosphere, while the reverse is observed for the lines of ionised elements (with the exception of the H and K lines). These facts would correspond to a higher mean temperature in the facula than in the adjoining photosphere.

We have said above that the determination of the effective temperatures of faculae (from the facula-photosphere contrast) is hindered by the peculiarities of the law of temperature distribution within the faculae. Hence we may attempt to determine $(T_e)_f$ as follows. We assume that we have determined the difference of equivalent widths $\Delta W_\lambda = (W_\lambda)_f - (W_\lambda)_\odot$ for a number of absorption lines. We further assume that the dependence of W_λ on the spectral class of the star, and therefore (see Table 11) on the effective temperature, is known for these lines. Then, supposing that the physical conditions in a facula are similar to those in the atmosphere of a star whose effective temperature is higher than that of the Sun while g has the same value, we can determine the difference $\Delta T_e = (T_e)_f - (T_e)_\odot$ from the values found for ΔW_λ. By this means ΔT_e is found to be about a hundred degrees. If we take the "filamentary" structure of faculae into account, ΔT_e in the "filaments" is still greater.

It is extremely important that further work should be done on the application of this method, since, for example, a difference in ΔT_e for neutral and ionised atoms might indicate a deviation from the conditions of thermodynamic equilibrium in the faculae.

The obtaining of a reliable estimate of the electron pressure in faculae is made less easy by various anomalies in the behaviour of ionised elements (especially Ca II). Present data would appear to indicate that p_e in the faculae should not differ markedly from p_e in the adjoining photosphere.

The subject of the anomalous behaviour of Ca II lines in the spectra of faculae is not yet altogether clear. It is quite possible that we have here the same agency which we have suggested for the "metallic-line" stars (see Chapter 15). There are a number of indications that there is a large energy excess in the far ultra-violet region of the spectra of faculae. This must lead (see Chapter 15) to anomalously strong ionisation of the Ca II atoms and a diminution in their number.

We shall briefly consider the behaviour of hydrogen lines in the spectra of faculae. The contours of the H_α line in these spectra and in those of other features observed on the Sun are shown in Fig. 51. The continuous contour is the observed contour of the H_α line in the spectrum of a facula. However, in the segment *ab* the H_α line belonging to the facula proper should lie below the line *ab*, since *chromospheric flocculi* are situated above faculae, and these flocculi make a contribution of radiation in the central parts of H_α (see Chapter 21). The true contour of the H_α line in the spectrum of a facula, with the spurious radiation removed*, should go roughly along the line *bc*, which is shown dotted. The broadening of the H_α line contour in the spectrum of a facula, compared with the same line in the spectrum of the adjoining parts of the photosphere, is caused (it is now thought) by the increased excitation of the hydrogen atoms in the faculae. In consequence of this, the number N_2 of hydrogen atoms in the second quantum state is increased, and the line is thereby broadened. We should expect a decrease in the central residual intensity in the facula, compared with the photosphere, again because of the increase of N_2 in the facula. Moreover, this decrease of r_{ν_0} is often observed in the H_β line†, where the

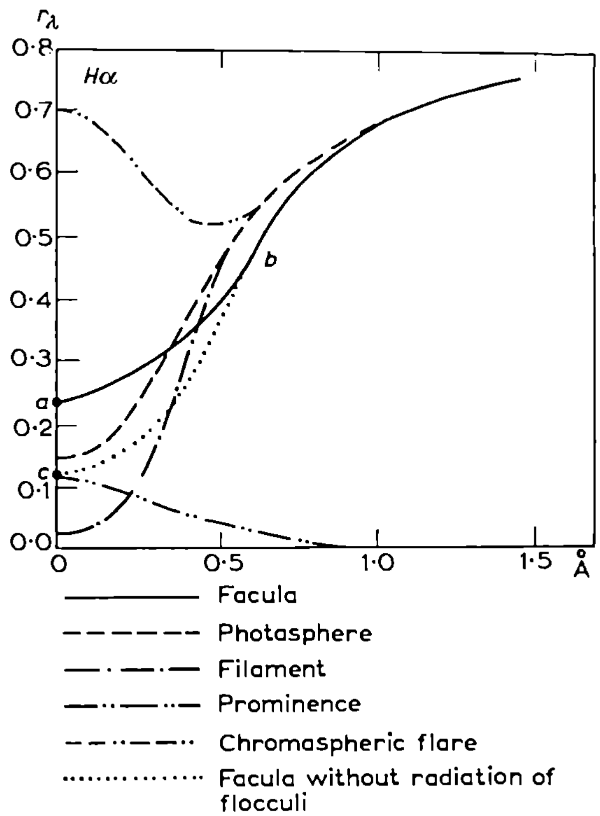

FIG. 51

* That is, the H_α line contour in the spectrum of a facula in the case where there are no chromospheric flocculi.

† This is often seen in spectroheliograms taken in the H_β line. In such cases the facular area is darker in H_β than the neighbouring parts of the phótosphere.

radiation belonging to the flocculi is much weaker than in H_α, and consequently does not noticeably distort the H_β line contour belonging to the facula proper. The same is true to an even greater extent of the higher terms of the Balmer series.

If the H_α line were broadened because of the pressure effect (possible increase in n_e) and not because of an increase in the number of atoms in the second state, we should expect a considerable broadening of the H_β, H_γ, . . ., lines, where the pressure effect is much greater than for H_α. However, in the faculae the broadening of these lines is considerably less than that of the H_α line.

The fact that hydrogen atoms are more strongly excited in faculae than in the adjoining photosphere does not contradict the approximate equality of $(T_{ex})_f$ and $(T_{ex})_\odot$ for lines of metals, found from observation. The hydrogen atoms which produce the Balmer series, unlike the metal atoms, have an excitation energy of 10·16 eV. Here even a small increase in T_{ex} can lead to a considerable increase in the ratio N_2/N_1. Furthermore, it is quite possible (as we have already said) that there are, in the faculae, deviations from thermodynamic equilibrium which, as usual, increase with increasing excitation potential.

Finally, observation shows that the temperature distribution in different faculae is very different. Thus, for example, in the spectrum of a very bright facula observed on 13 September 1950, the contours of the lines H_α, H_β, H_γ, and H_δ, including the wings, were less deep than the corresponding contours in the spectrum of the adjoining photosphere [90]. This fact cannot be explained by an increased ionisation of hydrogen, since this would lead to an absurdly high electron pressure in the facula, of the order of 10,000 bars. We must therefore, apparently, suppose that the smaller number N_2 of hydrogen atoms was due to a lower excitation temperature for hydrogen in the facula concerned.

Chapter 20. Prominences

1. The classification of prominences. In studying prominences we meet with a collection of phenomena which are extremely complex, and rich in physical significance. The application of cinematography to the study of solar phenomena has made possible a *provisional* classification of prominences. This classification must be regarded only as a division into various forms of prominence differing from one another, with an indication of the possible connection between these forms.

The features on which this classification of prominences is based are:

(1) The external form and structure of the prominences.

(2) The nature of the motion of matter in the prominences.

(3) The relation between prominences and sunspots.

(4) The place where the prominences are formed (chromosphere or corona).

We shall not explain in detail the present-day classification of prominences, but shall confine ourselves to some very brief remarks*.

Quiescent prominences (often resembling haystacks in appearance) may remain almost unchanged during many hours. One of the best criteria distinguishing quiescent prominences is the absence in them of external currents or jets, which are one of the main distinguishing features of *active prominences.*

Ordinary active prominences are the kind most frequently found on the Sun. They are observed *in all latitudes*, and this distinguishes them from *active sunspot prominences* (see below). The most characteristic property of ordinary active prominences is the presence of currents of moving matter, which emerge from the outer parts of the "trunks" of these prominences and proceed downwards to the chromosphere, into which they are drawn, as it were. The place where these currents flow into the chromosphere, which lies at some distance from the "trunk", is called the **centre of attraction.** Judging by the observations at present existing of the process of inflow of matter, the centres of attraction are relatively small areas in the chromosphere, which are not distinguished by any external features. There may be several centres of attraction for one prominence. The currents of matter which move to the centre of attraction usually consist of separate *jets* (streamers) and *knots*. The length of a jet varies from a few thousand kilometres to several hundred thousand kilometres in very active prominences. Large wide jets may rapidly dissipate the prominence.

Furthermore, if the centre of attraction is extremely strong, the whole of an active prominence may be torn off from the trunk which joins it to the chromosphere, raised to a height in a strongly curved trajectory and then, moving down, be drawn into the centre of attraction (such a prominence is said to be quasi-eruptive).

What are called **interacting prominences** also belong to the class of active prominences. Here the matter flows across, in the form of jets and knots, from one prominence to another. Cases are also observed where these jets and knots move in both directions; there is here, consequently, an exchange of matter, between the two prominences.

Eruptive prominences are those which rise relatively quickly to great heights above the Sun's surface, and then disappear. The maximum heights reached by these prominences lie chiefly between 100,000 and 500,000 km. At greater heights, the material in eruptive prominences

* We shall use the terminology employed in E. PETTIT's classification, which is one of the most complete.

is dispersed and ceases to be visible. These prominences are formed from quiescent or active prominences. No eruptive prominence seems to be known which rose directly from the chromosphere.

The process of the formation of an eruptive prominence from an active one* can be regarded as an explosion, although, of course, this word refers to the external appearance of the phenomenon, and not to its nature. As a result of the explosion, the prominence, which was until then an active prominence, expands, and after rising disappears against the background of the corona. During the whole process of expansion, the prominence sends out currents of luminous matter downwards. These currents, which are fairly numerous at first, are directed towards the earlier centres of attraction and towards newly-formed centres.

Eruptive prominences may apparently be divided into two subclasses. The first class contains those in which a relatively isolated mass of gas rises, while downward currents leave it. The second class consists of eruptive prominences having the form of an arc whose vertex rises while the whole arc expands. Here the downward currents are directed along the two sides of the arc.

In conclusion, it must be noted that the following circumstance is extremely important. The appearance of an eruptive prominence (i. e. the explosion of an active prominence) usually takes place in regions where neither photospheric nor chromospheric disturbances are observed.

Sunspot prominences are divided into a series of greatly differing sub-classes. The common feature of all these is that *they are related to sunspots*. Prominences of this class are never observed without accompanying spots, though spots may frequently be observed without prominences. We shall here refer to the most typical of the sunspot prominences.

The kind of sunspot prominences most often observed are what might be called **coronal sunspot prominences.** Here the luminous matter forming the prominence appears in the corona above the spot. What happens is that, above a group of sunspots, convergent currents of matter (jets) or detached parts of currents appear and all move towards the spots in straight or curved trajectories. In its external appearance, such a prominence resembles an opened fan. Knots of concentrated luminous matter are often observed in the upper parts of these currents; these knots are generally formed in the coronal region also, and, remaining at rest, serve as "reservoirs" for the downward flowing jets. It is of interest to note that the brightness of the jets increases as they approach the chromosphere.

* More exactly, from a quiescent one, since the active prominence is formed only a relatively short time before the "eruption".

In the case of the most active groups of sunspots, the matter in the knots flows downward, not in one direction, but in several, and closed loops are thus formed. Such prominences resemble fountains in external appearance. In rare cases the loops may arise from the chromosphere also.

Coronal sunspot prominences are sometimes accompanied on one or both sides by active prominences similar to those discussed above, with which they have many features in common. In this case the currents of luminous matter flow into the region of the spots.

A very interesting class of sunspot prominences is formed by those which are called **surges**. Here the matter is ejected from the chromosphere at speeds which sometimes exceed 500 km/sec, and then returns *along almost the same path back to the chromosphere.* Surge prominences may either be small condensations, rising to several thousands of kilometres and subsequently returning to the chromosphere, or take the form of huge branches of luminous matter with a height of over 100,000 km. In the latter case, practically all the luminous matter in the prominence returns to the chromosphere. Occasionally the upper parts only are detached from it.

The lifetime of a surge prominence is very short: from 10 to 20 minutes for small ones, and a little over an hour for large ones. It is very interesting that the ejection of surge prominences can take place at any angle to the Sun's surface.

The ejection of small aggregates of chromospheric matter (often at large angles to the vertical) is observed in fairly active spot groups; like the eruptive prominences, they do not return to the chromosphere.

Besides the forms of sunspot prominence mentioned, there are others which we shall not discuss here.

Tornado-type prominences are like vertical spirals or coiled-up ropes. In this case we are evidently concerned with the rotation of the prominence and in general with phenomena of a *vortex* nature. There are indications that rotation occurs in prominences of other types also.

Finally, we must briefly mention the prominences whose formation is connected with the coronal region. We have already spoken of coronal prominences of the sunspot type. A similar case is found with ordinary active prominences. As the number of streams of luminous matter emerging from such a prominence increases (this indicates that the centre of attraction is playing an increasing part), long luminous jets begin to pass now and again from the region of the corona to the centre of attraction; they have a small curvature and a velocity of the order of 100 to 200 km/sec. They may appear at heights of about 150,000 km and usually move with constant velocity along their whole path to the centre of attraction. Observations show that these prominences may

appear and enter the chromosphere even when no other prominence is observed; this shows that centres of attraction can exist on the surface of the Sun independently of prominences.

The **coronal clouds** are a very interesting form of coronal sunspot prominence. Their development takes place as follows. Above a group of sunspots, bright points are suddenly formed. New points subsequently appear, whereupon all the points coalesce into a cloud, which, as it were, hangs in the coronal region, while streams of matter in the form of jets begin to move down from it to the sunspots.

In concluding our description of the various classes of prominences, we must again emphasise that the present classifications have as yet no physical basis. However, some classification of prominences, if only a very provisional one, is indispensable, since, as we have seen, there is a very great difference in the forms of various types of prominences, and to interpret the physical properties of each type a separate approach must be made, though some regularities may, of course, be more general in character.

It must be noted that the actual development of prominences is much more complicated than can be realised from a single simple classification. Sometimes it is very difficult even to assign prominences to one or another class. Finally, it is necessary to take account also of the interconnection between prominences of various classes. For example, eruptive prominences may appear as an outcome of the development of active prominences, and so on.

The classification which we have considered is based on observations of prominences at the limb of the solar disc. Here the prominences are observed in emission. If a prominence lies between the observer and the photosphere and is projected on the Sun's disc, it absorbs the latter's radiation (in lines where the optical thickness of the prominence is large) as a comparatively dark formation called a **filament**. In the majority of cases, the filaments are of small width and have a very elongated form. Their length (estimated in projection) is sometimes a considerable fraction of the Sun's diameter.

As they approach the limb of the disc in consequence of the Sun's rotation, the observer usually sees the filaments transformed into prominences; however, there is not a complete correspondence between the geometrical characteristics of filaments and of the prominences which arise from them. The filaments are apparently seen as a result of the absorption of photospheric radiation, in various frequencies, by the densest and most extended parts of the prominences.

From the phenomena of "transformation" of filaments into prominences and *vice versa*, it has been established that, in the majority of cases, filaments are formed by quiescent prominences. It is of interest

to remark that these filaments (and therefore quiescent prominences also) avoid regions occupied by sunspots. If a spot appears close to a filament, the latter rapidly vanishes. In general, however, filaments exist for a comparatively long time, measured in days or weeks.

Sometimes sunspot prominences too cause a marked attenuation of the radiation of the solar disc in the line frequencies. However, in this case the dark formations on the disc are smaller than ordinary filaments. They have relatively short lifetimes and are different in form; they resemble small dark elongated areas of irregular shape, very dark small spots, etc. These formations are very often observed above *chromospheric flares*.

2. The motions of matter in prominences. Let us now consider the subject of the motions of the matter which forms prominences. A very great amount of information on this subject has been obtained by means of the cinematography of prominences, using narrow-band interference-polarisation filters. In the Soviet Union, such cinematography was first achieved by A. B. Severnyĭ and A. B. Gil'varg [133] at the Crimean astrophysical observatory of the USSR Academy of Sciences. By studying the material obtained, A. B. Severnyĭ [130] was able to define three basic types of motion in prominences.

The first type includes the motions observed when eruptive prominences break out. The second and most numerous type of motion contains, according to A. B. Severnyĭ, quite regular accelerated or retarded motions of individual knots, jets and clouds in prominences along curved trajectories. These motions are characteristic of sunspot prominences, active prominences, coronal clouds and surge prominences.

Finally, the third type of motion includes irregular motions of individual knots, clouds and jets (like waves), and also changes of shape and structure, which are very similar to the metamorphoses of terrestrial clouds, or smoke drifting on an uneven surface. Such metamorphoses, sometimes extremely slow and lasting for hours or days, are characteristic of quiescent prominences. In this case we are doubtless concerned with the appearance of the same turbulence which is found throughout the solar atmosphere.

It is clear that such a division of motions into types must form the starting-point for the construction of a future well-founded classification of prominences.

Let us now consider other regularities discovered in the motions of prominences. The use of cinematograph films has made it possible to ascertain that the motions of matter *down into the chromosphere* must be regarded as of much greater importance than was earlier supposed. As an example, we may point to coronal sunspot prominences, active prominences, etc. Among the numerous sub-divisions of sunspot prominences,

only a few appear in the chromosphere (for instance, surge prominences). The remainder are formed in the corona, whence the streams of luminous matter flow downwards. In general, contrary to the views previously held, it must be supposed that many classes of prominence are formed in the corona.

These circumstances are taken into account, in particular, in a new classification of prominences made by D. H. Menzel and J. W. Evans [85]. In this classification, prominences are divided into two main classes: (A) those formed in the corona, the matter in which moves predominantly downwards, and (B) those formed in the chromosphere, in which the matter moves upwards.

A careful examination of cinematograph films of the motion of prominences shows that, in a very large number of cases (especially for coronal sunspot prominences, active prominences and interacting prominences), the motion of the knots and jets of luminous matter takes place for long periods along certain quite definite curved trajectories. This suggests a field of some kind of force in the region occupied by the prominences.

A study of eruptive prominences has shown that, as a prominence moves outward from the Sun, its velocity increases, and the change in velocity takes place very quickly, almost *in jumps*, if we follow the leading edge of the jet or knot. Between the moments when the velocity suddenly increases, it remains *constant*, or varies very slowly. (In reality, of course, the velocity does not change instantaneously, but during a very short interval of time.) The time intervals between these velocity jumps are measured sometimes in minutes, sometimes in hours.

A. B. Severnyĭ [128] has discovered that, near the moment when the velocity suddenly changes, the brightness of the knots in the prominences increases, and, at the moment when the velocity changes sharply, reaches a maximum, though these outbursts of brightness take place, of course, simultaneously with a general decrease in the mean brightness of the prominence as it rises.

In Fig. 52 we give graphs taken from A. B. Severnyĭ's work. These illustrate the regularities mentioned above concerning the changes of velocity and brightness of the knots in eruptive prominences.

The result just discussed must also be related to the circumstance that an eruptive prominence becomes brighter just before its upward flight.

The following remark should be made in concluding the subject of the changes in velocity in the motion of matter in eruptive prominences. We have shown above that the changes in velocity in the motions of an eruptive prominence may occur very rapidly, almost in jumps (Pettit's first law). However, in studying such a jump, very great care must be taken, since in measuring the velocity we must rely on a determination

of the position of the individual luminous knots. There may, however, be some ambiguity in the identification of knots at different instants. In particular, L. Larmore [78], in his work on four prominences (including a large eruptive prominence), finds no case of a sudden change in velocity. This question is extremely important in the physics of prominences, and requires further study.

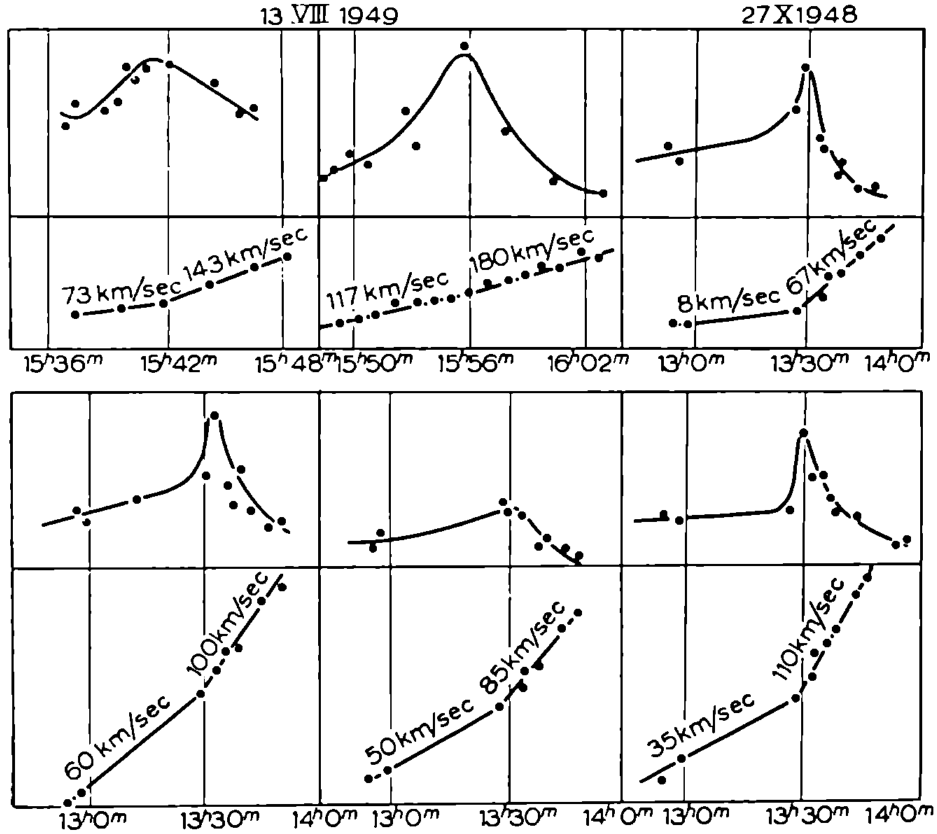

Fig. 52

Observations have shown that the atoms of various elements in prominences move with practically the same velocity. In other words, there is no separation of the elements in prominences. This is certainly a result of the presence in prominences of sufficient forces of internal resistance to motion. If, for example, calcium atoms are given an acceleration greater than that of other elements, the Ca atoms will have to "push through" a crowd of other atoms, i.e. the latter will exert a resistance to the moving Ca atoms. It is obvious that this resistance will be greater, the greater the density of the medium. Since the various atoms in prominences move with the same velocity, we must suppose that the density of prominences is sufficiently great to prevent a separation of the elements. The same follows from theoretical calculations also (see, e.g., W. H. McCrea's calculations [168, p. 677]).

3. **The spectra of prominences.** Having acquainted ourselves with the general properties of prominences, let us now consider and interpret their spectra. First of all, it must be remarked that the spectrum of a prominence projected on the sky is an *emission spectrum*. It consists of bright lines, a faint recombination continuum beyond the limit of the Balmer series, and an even fainter continuum due to the scattering of solar radiation by the free electrons contained in prominences. These characteristics of the spectra of prominences are related to the fact that the optical thickness of a prominence in the frequencies of the continuous spectrum is considerably less than unity (as follows from observation).

We shall explain what we have said about the bright-line spectra of prominences by considering the case where the optical thickness of a prominence in the line frequencies is also less than unity. In this case (since $\tau_\nu < 1$ in all frequencies), the intensity of radiation emerging from the prominence is proportional to the emission coefficient j_ν in all frequencies.

On the other hand, the emission coefficient is large in the frequencies where the absorption coefficient is large [see, e.g., formulae (3.10) and (9.5)]. And since, in the line frequencies, the absorption coefficient is greater than in the neighbouring parts of the continuous spectrum, the emission coefficient also will be greater in the line frequencies. Hence, in projection on the background of the sky, the prominences must in fact give a bright-line spectrum. Here it is clear that the result in question does not depend at all on the simplifying assumption which we have made, that the prominence is transparent in the frequencies of the absorption lines.

If, however, the prominence is projected on the solar disc, the absorption lines of the solar spectrum at the corresponding part of the disc must be deepened by the additional absorption of light in the prominence. This is indeed observed for some lines in the spectra of filaments, namely those lines where the optical thickness of the prominence is comparable with or greater than unity.

The bright-line spectra of quiescent prominences, observed *outside eclipse*, consist chiefly of the first few lines of the Balmer series of hydrogen, the H and K lines of Ca II, the D_3 line of neutral helium ($\lambda = 5876$ Å) and some fainter lines. The spectra of prominences obtained at the time of a *total solar eclipse*, when the effect of the daylight sky is reduced to a minimum, are much richer in lines and include several dozen emission lines. The majority of these are lines of hydrogen and of ionised metals. There are considerably fewer lines of neutral metals than of ionised metals, and they mainly correspond to fairly strong lines in the ordinary absorption spectrum of the Sun. Besides the lines mentioned, several lines of neutral helium (including D_3), and the ionised

helium line 4686 Å, are found in the spectra of quiescent prominences. Judging by the existing material, the spectra of eruptive prominences differ hardly at all from those of quiescent ones (in the relative intensities of the different lines).

Prominences related to spots are often called *metallic*. This name arose because their spectra, taken outside eclipse, contain, besides the lines of H, Ca and He, many bright lines of metals, such as Sr, Mg, Na, Ti, Fe, etc. These lines are sometimes so bright that the image of the prominence can be seen in their light even when the slit of the spectroheliosoope or prominence spectroscope is open. In some cases, a fairly intense continuous spectrum has also been observed in prominences of this class.

However, there is apparently no essential difference between the spectra of quiescent prominences and those of sunspot prominences. A comparison of spectra (obtained at the time of an eclipse) shows that the former exhibit mainly lines which have intensities, in the scale commonly used for estimating the intensities of chromospheric lines, of more than 30, whilst in the metallic prominences fainter lines also appear, with intensities ≥ 15. Thus the spectral lines of quiescent and metallic sunspot prominences differ from each other chiefly by their intensities alone.

This difference may be due either to a difference in the degree of ionisation and excitation, or to a difference in the number of emitting atoms (but not to a difference in chemical composition). We should expect that the number of emitting atoms in sunspot prominences is greater than the corresponding number in quiescent prominences.

Many of the characteristic properties of prominence spectra, such as the presence in these spectra of He I and He II lines, the predominance of the lines of ionised metals compared with those of neutral metals, and so on, correspond to a higher state of excitation and ionisation of atoms than under the conditions of the photosphere and the "reversing layer" of the Sun. For example, A. Unsöld points out that the spectrum of an eruptive prominence in emission has exactly the same properties as an ordinary spectrum of a star of class cA. However, more definite conclusions must be based on a quantitative analysis of prominence spectra. This we shall now give.

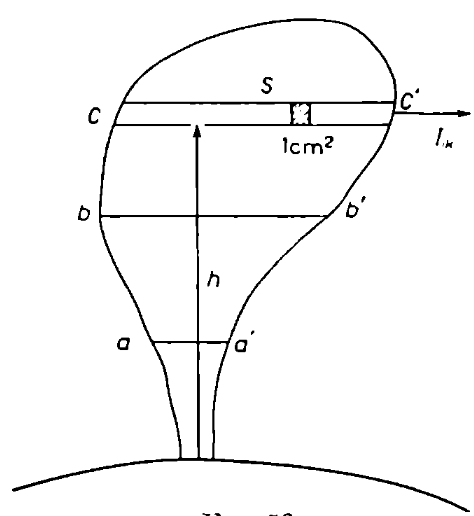

Fig. 53

4. Self-absorption in prominences. We draw a cylinder in the prominence, directed along the line of sight and having unit base area (Fig. 53). The length s of this cylinder, which is at a height h above the photo-

sphere, is equal to the thickness of the prominence in the given direction.

Here we shall be interested in the total energy emitted by the cylinder inside the line with transitions $k \to i$. Let I_{ik} be this total energy, emitted in 1 second by the cylinder towards the observer and referred to unit solid angle. It is evident that the quantity I_{ik} is also the total intensity of radiation in the given line and in the given direction.

Let the number of emitting atoms in the cylinder be N_k. We assume that the optical thickness of the prominence in the direction considered, inside the whole line, is considerably less than unity. In other words, we assume that we can neglect self-absorption inside the prominence (in the given direction). Then we have for I_{ik}

$$I_{ik} = \frac{N_k A_{ki}}{4\pi} h\nu_{ik} \,. \tag{20.1}$$

since we can neglect processes of stimulated emission, for not very small $h\nu_{ik}$, under the conditions existing in the Sun. If we have determined the value of I_{ik} in some way from observation*, the value of N_k can be determined from (20.1). However, in many very important cases, it is necessary to take account of self-absorption. Let us briefly consider this subject.

Replacing the quantity $\varkappa_\nu$ in equation (2.26) by σ_ν, and denoting the element of optical thickness along the direction cc' by $d\tau_\nu = \sigma_\nu \varrho \, ds$, we obtain

$$dI_\nu/d\tau_\nu = -I_\nu + j_\nu/\sigma_\nu \,. \tag{20.2}$$

The solution of this equation, which is found in the same way as (3.36) and (3.37), gives

$$I_\nu = \int_0^{\tau_\nu(s)} \frac{j_\nu}{\sigma_\nu} e^{-\tau_\nu} \, d\tau_\nu \,, \tag{20.3}$$

where $\tau_\nu(s)$ is the optical length of the whole cylinder cc'. The optical length can be measured either from the point c or from c'. Integrating (20.3) over all frequencies in the emission line, we clearly obtain I_{ik}:

$$I_{ik} = \int I_\nu \, d\nu = \int \left\{ \int_0^{\tau_\nu(s)} \frac{j_\nu}{\sigma_\nu} e^{-\tau_\nu} \, d\tau_\nu \right\} d\nu \,. \tag{20.4}$$

* Observers usually express the energy emitted by prominences in terms of that emitted by the centre or limb of the solar disc.

If self-absorption is absent $[\tau_\nu(s) \ll 1]$, the expression (20.4) takes the form

$$I_{ik} = \int \left\{ \int_0^s j_\nu \, \varrho \, ds \right\} d\nu \, . \tag{20.5}$$

The computation of the integral on the right-hand side of (20.4) presents great difficulties in the general case. We shall therefore simplify the problem somewhat. Assuming that the physical conditions inside a prominence depend only on the height h above the photosphere, we can regard the quantities j_ν and σ_ν as constant along the path of integration (the cylinder cc'). We can take j_ν/σ_ν outside the inner integral sign in (20.4). Further, instead of $\tau_\nu(s)$ we can write, using (5.20), (5.22) and the constancy of σ_ν ,

$$\tau_\nu(s) = \int_0^s \sigma_\nu \, \varrho \, ds = m \int_0^s \sigma_\nu \, n_i \, ds = s_\nu \, N_i \, , \tag{20.6}$$

where s_ν is the absorption coefficient referred to one atom, and N_i is the number of absorbing atoms in the cylinder cc'. Thus we have, instead of (20.4),

$$I_{ik} = \int \frac{j_\nu}{\sigma_\nu} \left(1 - e^{-s_\nu N_i}\right) d\nu \, . \tag{20.7}$$

We now introduce the notation

$$P_\nu = j_\nu/\sigma_\nu . \tag{20.8}$$

To ascertain the processes by which the value of P_ν is determined is the greatest difficulty in the problem of self-absorption. If local thermodynamic equilibrium existed in the prominence (in fact it does not), then, by (3.10), we should have $P_\nu \equiv B_\nu$. If the prominences simply scatter the solar radiation which is incident on them from below, then, according to (9.5), $P_\nu = J_\nu$. Finally, if the emission in the prominences is a result of fluorescence or the excitation of atoms by electron collisions, P_ν is altogether different. However, in all these cases the quantity P_ν must vary relatively little with frequency inside the line (considerably less than the expression in parentheses in the integrand). Hence (20.7) can be rewritten

$$I_{ik} = \bar{P}_\nu \int \left(1 - e^{-s_\nu N_i}\right) d\nu, \tag{20.9}$$

where $\bar{P}_\nu$ is the mean value of P_ν.

The appropriate investigations show that, even in the centres of strong lines, the optical thickness of prominences does not exceed 10 (or at most a few times this). This means (see Fig. 25) that in the damping

region (where s_ν is several orders of magnitude less than s_{ν_0}) the optical thickness $s_\nu N_i \ll 1$. Consequently, we can take for s_ν simply the Doppler coefficient given by formula (11.34), where $\Delta\nu_D$ is determined by both thermal and turbulent motions.

If the value of P_ν in any direction (e.g. aa', bb', cc', etc.) is the same, the quantity I_{ik} will be determined only by the value of the integral on the right-hand side of (20.9), and consequently will depend only on N_i and on the parameter v_0 in (11.29), so that we shall have some *curve of growth* (that is, a relation between I_{ik} and N_i), but *for emission lines*. If, also, the function P_ν is constant for all prominences of the class concerned, then such a curve of growth can in principle be constructed from many prominences. A family of such curves of growth constructed for various values of the parameter v_0 may form the starting-point both for a comparison of theory and observation and for a determination of the quantities N_i and v_0. The values of the above integral have been calculated, for example, by E. F. M. VAN DER HELD [169].

Assuming for the moment that this hypothesis about the constancy of P_ν is correct, we can decide about the presence of self-absorption as follows. We select two lines: one for which self-absorption is certainly absent, say the D_3 line of He I, and the other for which we should expect considerable self-absorption, say the H_α line. Having measured the values of I_{ik} for both these lines at various points in the prominence (or even in different prominences), we construct graphs, placing $(\log_{10} I_{ik})_{He}$ on the axis of abscissae and $(\log_{10} I_{ik})_{H_\alpha}$ on the axis of ordinates. If now the conditions of excitation and ionisation are the same throughout the prominence (or all the prominences), and self-absorption is absent, then there should be a proportionality between $(I_{ik})_{H_\alpha}$ and $(I_{ik})_{He}$, as follows from (20.1). However, if self-absorption is present, it is quite clear that $(I_{ik})_{H_\alpha}$ is *diminished* where it is large. For, the larger I_{ik}, the greater N_k, and consequently (for P_ν constant) the greater N_i, i.e. the greater the self-absorption. As a result, we obtain a graph like Fig. 54. The quantity Δ is a measure of the self-absorption. From such investigations it has been discovered that the H and K lines of Ca II usually have strong self-absorption, the H_α line moderate self-absorption, and the D_3 helium line, like all the fainter lines, practically no self-absorption.

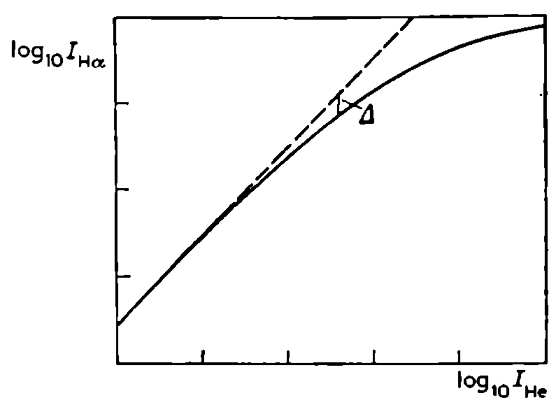

FIG. 54

The method just described has the defect (which is revealed by the scatter of the points about curves like Fig. 54) that in using it we postu-

late the existence of a universally valid curve of growth, i.e. the constancy of $\bar{P}_\nu$. However, it is quite clear that, besides any other factor, the transition from an optically thin to an optically thick layer will itself cause a change in $\bar{P}_\nu$. This is particularly apparent in the case where the value of I_{ik} is determined by processes of scattering of solar radiation. It is evident that, for $\tau_{\nu_0} > 1$, the value of J_ν will be different in the outer and inner parts of prominences. The situation is similar for other mechanisms of the excitation of atoms.

For the H and K lines, according to V. P. Vyazanitsyn [175], this difficulty can be partly removed as follows. Since the H and K lines begin from the same level and lie in the same part of the spectrum, the excitation conditions for these lines must be practically the same, so that we can put $(\bar{P}_\nu)_{\rm H} \approx (\bar{P}_\nu)_{\rm K}$. Thus the effect of self-absorption on the ratio $(I_{ik})_{\rm H}/(I_{ik})_{\rm K}$ must be determined by the value of N_i only. Since the oscillator strength for the K line is twice that for the H line, the self-absorption in the K line will be stronger than in the H line, and this effect will increase with N_i, i.e. with the total brightness of the prominence in the H and K lines. By means of (20.9), assuming, as shown above, that $(\bar{P}_\nu)_{\rm H} = (\bar{P}_\nu)_{\rm K}$, we can obtain the theoretical relation between N_i and the ratio $(I_{ik})_{\rm H}/(I_{ik})_{\rm K}$. The quantity v_0 which appears in the expression for the absorption coefficient can be determined from the half-width of the H and K lines. Using this method (with some modifications), V. P. Vyazanitsyn has obtained, from a large amount of material, a mean value of $N_i = N_1 = 2{\cdot}4 \times 10^{13}$ Ca II atoms for all the prominences which he studied. If we take the mean thickness of quiescent prominences as approximately 10,000 km, then 1 cm^3 of a prominence must contain about $2{\cdot}4 \times 10^4$ atoms of Ca II. From this material it was also found that self-absorption diminishes the brightness of the H line of Ca II by an average factor of three.

In another paper, V. P. Vyazanitsyn [176] has studied the rare phenomenon of the partial screening of one prominence by another. Here the prominence closer to the observer was represented by both emission lines and absorption lines.

5. **The excitation of atoms in prominences.** The principal problem in the theory of self-absorption is to establish the form of the function P_ν, i.e. to ascertain the excitation and ionisation conditions in prominences. However, it is quite clear that this is one of the main problems in the whole theory of prominences. In some cases it can be solved by means of formula (20.9). Let us assume, for instance, that the value of I_{ik} is determined by ordinary scattering of solar radiation in the prominence. We can then write $P_\nu = J_\nu$, where $J_\nu = \int I_\nu \, d\omega/4\pi = \bar{I}_\nu W$; W is the radiation dilution factor and $\bar{I}_\nu$ is the mean intensity of radiation inside

the prominence. If now we have determined (e.g. as shown above) the values of N_i and v_0, the integral on the right-hand side of (20.9) can be easily computed. Knowing the absolute value of I_{ik} from observation, we can determine $\bar{J}_\nu$ in accordance with (20.9). The value of $\bar{I}_\nu$ thus found should correspond closely to the mean intensity of solar radiation incident on the prominence. This is quite evident for small τ_ν. For large τ_ν, however, the radiation I_ν emerges, according to (20.3), from the exterior layers of the prominences where τ_ν is small (because of the exponential $e^{-\tau_\nu}$), i.e. here also the value of $\bar{I}_\nu$ should approximately correspond to the mean intensity of solar radiation. Calculations performed by V. P. Vyazanitsyn in the two papers mentioned have shown that, for the H and K lines, the value of $\bar{J}_\nu$ found in the manner described is in fact close to the value $J_\nu = \int I_\nu \, d\omega/4\pi$ for solar radiation in the centres of the H and K lines. Consequently, the radiation of prominences in the H and K lines is largely due to the scattering of solar radiation (in this case, resonance scattering). This result is related to the facts that the transition coefficients A_{ki} and B_{ki} for the H and K lines are large, and that the density of solar radiation in the region of the spectrum concerned is fairly high.

The same conclusion appears to be true also for many other lines (chiefly of metals) observed in the spectra of prominences. However, in a number of cases there is reason to suppose that the mechanism which makes prominences luminous mainly involves recombinations, accompanied by subsequent downward cascade transitions of the electron. For example, the intensities of emission lines of hydrogen and, in particular, of helium are anomalously high in comparison with the intensities of most other lines (chiefly those of metals with small excitation potentials).

By studying this problem we can obtain very valuable information on the nature of the ionising and exciting radiation from an examination of the distribution of the emitting atoms among the levels. Strictly speaking, we should solve this problem by starting from the corresponding *equations of the steady state*. Each of these equations asserts that the number of electrons arriving at the kth level of the r times ionised atom in 1 second by all possible paths is equal to the number of electrons leaving this level in 1 second, again by all possible paths. The two numbers are referred, of course, to the same volume, say 1 cm^3. (It is evident that these equations are valid in the case where the brightness of the prominence does not vary with time, or varies only slightly.)

Let us write down these equations. Let ϱ_{ik} be the density of radiation corresponding to the transition $i \to k$ or $k \to i$. We denote by C_{kf} the probability of photo-ionisation from the kth level. It can be computed from (14.20). We denote by C_{fk} the probability of photo-recombination to the kth level. According to (8.25), it is

$$C_{fk} = 4\,\pi\left(\frac{m_e}{2\,\pi\,k\,T}\right)^{3/2}\int\limits_0^\infty (\beta_\nu)_k\left(1+\frac{c^3}{8\,\pi\,h\,\nu^3}\,\varrho_\nu\right)e^{-m_e v^2/2kT}\,v^3\,dv. \quad (20.10)$$

Then the equation of the steady state for the first level, with $k = 1$, has the form

$$n_{r+1}\,n_e\,C_{f1} + \sum_{s=2}^{\infty} n_{r,s}\,(A_{s1} + \varrho_{1s}\,B_{s1}) = n_{r,1}\sum_{s=2}^{\infty}\varrho_{1s}\,B_{1s} + n_{r,1}\,C_{1f}\,. \quad (20.11)$$

The first term on the left-hand side gives the number of recombinations to the first level; the second gives the number of transitions to this level from higher discrete levels. The first term on the right-hand side gives the number of transitions from the first level to higher discrete levels. The second term gives the number of photo-ionisation processes from the first level.

The equation of equilibrium for the second level has the form

$$n_{r+1}\,n_e\,C_{f2} + \sum_{s=3}^{\infty} n_{r,s}\,(A_{s2} + \varrho_{2s}\,B_{s2}) + n_{r,1}\;\varrho_{12}\,B_{12}$$
$$= n_{r,2}\,\{\sum_{s=3}^{\infty}\varrho_{2s}\,B_{2s} + A_{21} + \varrho_{12}\,B_{21} + C_{2f}\}. \quad (20.12)$$

For the third level, we have

$$n_{r+1}\,n_e\,C_{f3} + \sum_{s=4}^{\infty} n_{r,s}\,(A_{s3} + \varrho_{3s}\,B_{s3}) + \sum_{s=1}^{2} n_{r,s}\;\varrho_{s3}\,B_{s3}$$
$$= n_{r,3}\,\{\sum_{s=4}^{\infty}\varrho_{3s}\,B_{3s} + \sum_{s=1}^{2}(A_{3s} + \varrho_{s3}\,B_{3s}) + C_{3f}\}. \quad (20.13)$$

The equations for the other levels are written down similarly. It is clear that there must be an infinite number of them. However, in practice, we restrict ourselves to a finite number of equations, since, when these are correctly chosen, the remaining equations have an insignificant effect on the result.

In constructing the equations of the steady state, we have neglected the excitation of atoms by collisions, because of the low density of matter in prominences. (The kinetic temperature of prominences, judging from various data, is about 6000° to 15,000°.)

To find the distribution of atoms among states, we need to know the quantities ϱ_{ik} and C_{kf}.

Having taken a series of values for the temperature of the radiation (excitation temperature) T_1 up to the limit of the Lyman series, and of the ionisation temperature T_2 beyond this limit, V. P. VYAZANITSYN

[175] was able to obtain a fairly satisfactory explanation of the observed ratios $I_{H_\beta} : I_{H_\gamma} : I_{H_\delta} : I_{H_\varepsilon}$. It was found that $T'_1 = 5670°$ and $T'_2 = 6650°$. Thus the excitation temperature is here close to the effective temperature of the Sun. On the other hand, the value $T'_2 = 6650°$ indicates the presence of a definite excess in the intensity of solar radiation beyond the limit of the Lyman series*. This leads to an excess ionisation of hydrogen atoms, and thus to an increase in the number of recombinations, i. e. an increase in the brightness of the corresponding hydrogen emission lines.

It must be mentioned, however, that recombination emission seems to be important only in the higher terms of the Balmer series. It follows from recent investigations (G. S. IVANOV-KHOLODNYĬ, P. TEN BRUGGENCATE) that the occupation of the third and part of the fourth level of the hydrogen atom corresponds to a relatively low excitation temperature, of the order of 4000°. This may be taken to indicate that the excitation in the frequencies of H_β and (in particular) H_α can be effected even by solar radiation that is attenuated by the presence of the Fraunhofer absorption lines H_β and H_α and by virtue of the inequality $W \ll 1$, where W is given by formula (8.3).

In connection with this problem, we must make the following remark. Although the above way of solving equations of the type (20.11) to (20.13), etc., is fundamental and physically the most correct, it nevertheless involves a number of serious difficulties. The chief of these is that the density of the ionising and exciting radiation in the far ultra-violet region of the spectrum is unknown to us, while the presence of lines of He I and He II in the spectra of prominences indicates that we should expect many anomalies as regards the density of radiation.

In particular, the nature of the ionising radiation in prominences, both as regards temperature and as regards its source, may be markedly different from that of the exciting radiation. Thus, the ionisation of hydrogen and helium in prominences is apparently determined by hard *coronal radiation*. On the other hand, the excitation of hydrogen atoms from the first level is apparently determined by the *emission of the prominence* in the lines of the Lyman series (resulting from recombination processes with subsequent cascade transitions). It is true that direct chromospheric radiation in the Lyman α, β, etc., lines also falls on the prominences. However, the optical thickness of the prominence in these lines is very large (in the H_α line it is often comparable with unity), so that monochromatic chromospheric radiation penetrates into the prominence only in a very much attenuated form.

* On account of the high opacity of the solar gases beyond the limit $\lambda = 912$ Å, we should expect that the intensity of radiation here would correspond to the boundary temperature of the solar photosphere, which is approximately equal to 4500° or 5000°.

These results show that our information about the values of ϱ_{ik} and C_{kf} for high frequencies is as yet very uncertain. In general, the theoretical conclusions regarding transitions with high excitation energies are still very indefinite, the more so since the situation may differ for different classes of prominence. On account of this indefiniteness, various indirect methods are employed [besides the use of equations (20.11) to (20.13)]. One of these methods, which is very widely used, is the following.

Let us consider atoms with a large first excitation potential, such as hydrogen or helium atoms. The application of the formulae of thermal equilibrium to these atoms can, as we have shown, lead to large errors. On the other hand, there is reason to suppose that, for transitions with small energies, the use of these formulae is much more legitimate. We therefore consider transitions between the continuum and levels having a comparatively small binding energy $\chi_{r,k}$ (up to 5 eV). For these transitions we can use, on the above grounds, the formulae which are valid in the presence of thermodynamic equilibrium. We obtain one such formula by dividing (5.54) by (5.11) and taking account of (5.6):

$$n_{r,k} = n_{r+1}\, p_e\, \frac{g_{r,k}}{u_{r+1}}\, \frac{h^3}{2(2\pi m_e)^{3/2}(kT)^{5/2}}\, e^{\chi_{r,k}/kT}, \tag{20.14}$$

or, in logarithmic form,

$$\log_{10} n_{r,k} = \log_{10} n_{r+1} + \log_{10} p_e + \log_{10} \frac{g_{r,k}}{u_{r+1}} + \\ + \log_{10} \frac{h^3}{2(2\pi m_e)^{3/2} k^{5/2}} - \tfrac{5}{2}\log_{10} T + \frac{5040}{T}\chi_{r,k}, \tag{20.15}$$

where $\chi_{r,k}$ is now expressed in electron-volts.

If now we determine $N_{r,k}$ by means of (20.1) or (20.9), then, knowing p_e and T, we can find N_{r+1}, and so on.

Since formula (20.14) is used by many authors, and a large number of quantitative investigations have been performed by means of it, we shall examine its validity. It can be shown that formula (20.14) would hold completely rigorously if the following three conditions were satisfied: (1) the density of radiation incident on the atoms, in the long-wave part of the spectrum (where the values of $h\nu$ correspond to the small values of $\chi_{r,k}$ under consideration), is determined by Planck's formula; (2) the velocity distribution of electrons corresponds, for low energies, to Maxwell's formula; (3) transitions between the ground level (and the low levels in general), of the atom or ion concerned, and the high levels in which we are interested, are forbidden or do not play an important part.

The first condition is fulfilled with sufficient accuracy in prominences. The second condition also is satisfied (see Chapter 8). It is more difficult to decide how far the third condition is fulfilled. It holds, for example, for the levels of orthohelium. But the fact that the application of formula (20.14) leads in a number of other cases also to plausible and self-consistent results indicates that in these cases the number of transitions from the ground state to any excited level *does not exceed*, in order of magnitude, the number of transitions between excited levels and the number of recombinations to this excited level.

6. **The ionisation of atoms. The electron pressure.** Let us now briefly consider the question of the ionisation of atoms in prominences. Our first task here is to determine the mean value of p_e. We can do this by the following methods.

The first method is to determine p_e by means of formula (15.1), where m is the principal quantum number of the last distinguishable line in the Balmer or Paschen series in the prominence spectrum. For T, in passing from n_e to p_e, we can take a value of the order of 6000° to 15,000°; see below.

The second method of determining p_e is as follows. We shall see below that the hydrogen in prominences is ionised to a considerable extent, so that the number of free electrons in prominences is determined mainly by the ionisation of hydrogen, and not by the ionisation of metals. In this case $n_e \approx n_1$, where n_1 is the number of protons in 1cm³, and consequently, using equation (20.14), we obtain

$$n_{0,k} = n_e^2 \frac{g_{0,k}}{u_1} \frac{h^3}{2(2\pi m_e)^{3/2}(kT)^{3/2}} e^{\chi_{0,k}/kT}, \tag{20.16}$$

where $g_{0,k} = 2k^2$. Having determined the value of $n_{0,k}$ (for any Balmer or Paschen line, or even for several lines) by means of (20.1) and the known thickness of the prominence, we find the value of n_e by means of (20.16).

In accordance with what we have said above, we can take for T the temperature of the Sun's radiation in the red and infra-red parts of the spectrum, where it is close to 6000°. Since the exponential factor in (20.16) appears under a radical sign in the determination of n_e, we may suppose that the value found for n_e by this method will not differ much from the true value.

The third method of determining n_e was first used by V. P. Vyazanitsyn [175], and is based on the solution of equations (20.11) to (20.13). From these equations we find for hydrogen the ratio $n_{0,s}/n_e n_1 \approx$ $\approx n_{0,s}/n_e^2$, and knowing $n_{0,s}$ from observation we can find n_e.

Another method of determining n_e (and therefore the linear thickness of a prominence) has been developed by H. ZANSTRA and applied by various authors [182]. It is based on the following considerations. We have pointed out above that the faint continuous spectrum of prominences that is observed in the visible part of the spectrum is due to the scattering of solar radiation by free electrons in the prominences. This conclusion is based not only on a comparison of the coefficients j_ν for various emission processes, but also on the fact that, according to B. LYOT's observations, the continuous spectrum of prominences, at least in the range from 5950 to 6400 Å, is polarised to an extent of about 15 %.

Thus the *measured* intensity of the continuous spectrum of a prominence in the visible region should be equal to An_es, where s is the linear thickness of the prominence along the line of sight, and A is some constant determined by the electron scattering coefficient and the intensity of the solar radiation incident on the prominence. On the other hand, using formula (20.16), we can determine from observation the quantity $sn_{0,k}$, and therefore sn_e^2 if T is given. Since the two quantities sn_e and sn_e^2 are determined independently, we can then find s and n_e independently.

We can similarly use the ratio of the intensity of the continuous spectrum of a prominence in the visible region to its intensity just beyond the limit of the Balmer series. In the latter case, the radiation is due to recombination processes, and the intensity is again proportional to sn_e^2.

Various authors have applied these methods to determine n_e. The values of n_e found by these investigations lie between $n_e = 5 \times 10^9$ and $n_e \approx 10^{11}$ cm^{-3}. This variation is due partly to the difference in the methods employed, partly to our ignorance of the exact value of T_ε, and partly, perhaps, to actual differences in the physical conditions in different prominences.

The majority of the methods used to determine n_e presuppose a knowledge of the linear thickness of the prominence along the line of sight, so that the quantity n_e appears as a factor in the product sn_e^2. However, recent investigations by G. S. IVANOV-KHOLODNYĬ [62] show that the actual value of n_e in prominences (at least of some classes) may be considerably greater than the values just given. The chief observational facts which lead to this conclusion are as follows:

(1) We have mentioned above that the contours of emission lines in the spectra of prominences must be purely Doppler contours, if, of course, there is no self-absorption; this is true, in particular, for the high terms of the Balmer series. For these terms, therefore, the ratio found from observation between the half-width $\Delta\lambda$ of the line and the wavelength λ must, by (11.30), be the same for all these lines. On the

other hand, according to a great amount of material obtained by G. S. Ivanov-Kholodnyĭ, this ratio $\Delta\lambda/\lambda$ increases somewhat with the number of the line in the series. If we take this to be the result of Stark broadening, n_e is found to be of the order of 2×10^{12} to 2×10^{13} cm^{-3}.

(2) The highest line observed in the hydrogen emission from prominences is H_{31}. Fewer lines than this are usually observed, and the value of n_e resulting from this comparison is greater than 10^{12} cm^{-3}.

(3) The following effects are shown by helium lines:

(a) It is found that $(\Delta\lambda/\lambda)_{D_3} > (\Delta\lambda/\lambda)_{10,830}$, in accordance with the fact that the Stark effect should be greater for the higher levels than for the lower ones.

(b) The quantity $\Delta\lambda/\lambda$ for helium varies in correlation with its increase for hydrogen.

If now we take for sn_e^2 the value 10^{30} cm^{-5} which is usually obtained, and for n_e the value 10^{13} cm^{-3}, we find that $s = 10^4$ cm, i. e. the effective emitting layer of a prominence along the line of sight must be very thin. In other words, we come to imagine a prominence as an object consisting of numerous widely scattered *thin* luminous "threads" or "filaments".

It is clear that the above somewhat unexpected conclusions must be verified by further investigations. Nevertheless, the conception of prominences (at least of some classes) as consisting of numerous very thin threads is confirmed by various facts. Firstly, the striated, extremely thin, threadlike structure of many prominences was noticed even by the first investigators of these objects (A. Secchi, L. Respighi, C. A. Young, J. Fényi and others), who observed prominences visually at times when the images were very good*.

Furthermore, the idea that there are fine threads of luminous matter in prominences is in accordance with the fact that in many prominences there is observed (M. Waldmeier, M. N. Gnevyshev, R. S. Gnevysheva and others) the *coexistence* of two entirely different phases of matter, the prominence matter proper and coronal matter, whose kinetic temperature is of the order of 1,000,000° (see Chapter 22). This fact can, apparently, be explained only by supposing that between the two phases there exist some kind of "barriers", which seem to be of electromagnetic origin. The hotter coronal matter could, in fact, remain for a considerable time between the "threads". All these considerations require further development, of course.

* It is well known that the eye can detect considerably finer details and contrasts than can a photographic plate.

Let us now consider the problem of the origin of the anomalously high state of ionisation of the helium and hydrogen atoms in prominences. One of the most probable mechanisms of the ionisation of prominences, according to I. S. SHKLOVSKIĬ, is by the monochromatic ultra-violet radiation of the corona (see Chapter 22), i. e. the radiation of the corona in various lines lying in the far ultra-violet region of the spectrum. The chief of these are the coronal lines $\lambda = 776$ Å (Ne VIII) and $\lambda = 625$ Å (Mg X). The ionisation equation for this case is constructed similarly to equation (8.27), the quantities J_ν being different from zero only in the narrow regions of the coronal lines. In particular, according to (3.6) and (5.37), the number of photo-ionisations in 1 cm³ and 1 second caused by radiation of some given coronal line is, for hydrogen,

$$k_\nu' \, n_{0,1} \, \frac{c}{h\nu} \int \varrho_\nu \, \mathrm{d}\nu \, , \tag{20.17}$$

where $\int \varrho_\nu \, \mathrm{d}\nu$ is the integral of the density of monochromatic coronal radiation over the whole line, and k_ν' is the coefficient of photoelectric absorption from the first level, given by formula (5.24) with $n = 1$. The total number of photo-ionisations is determined by summing (20.17) over all the lines which cause ionisation. The number of photo-recombinations is determined as before by formula (8.25).

The use of such ionisation formulae does in fact explain the very high degree of ionisation of H and He in prominences. It appears that the ionising action of the coronal radiation mentioned is equivalent, for hydrogen, to Planck radiation with a temperature $T \approx 7000°$, which is close to the value of the "ionisation temperature" $T_2 \approx 6650°$, found by V. P. VYAZANITSYN. V. A. KRAT [72] arrives at the same quantitative conclusion.

Furthermore, such calculations show that the degree of ionisation x_{H} of hydrogen in prominences is of the order of 0·1 to 1·0. In this case the ionisation of hydrogen is indeed the principal source of free electrons. This circumstance (which we have already utilised) is in accordance with other results based on methods differing from the one used here. In Chapter 21 we shall consider one such method as applied to the chromosphere.

The deviations from the ionisation state corresponding to thermodynamic equilibrium can be estimated from formula (20.14), if there are, in the prominence spectrum, lines of two consecutive ionisation states of the element concerned. For prominences these are, of course, lines of neutral and singly ionised elements. No lines of doubly ionised elements are observed in the spectra of prominences.

From (20.14) we can write

$$\frac{n_{0,k}}{n_{1,k}} = \frac{n_1}{n_2}\frac{g_{0,k}}{g_{1,k}}\frac{u_2}{u_1} e^{(\chi_{0,k}-\chi_{1,k})/kT}, \tag{20.18}$$

where, of course, the suffix k may in general be different for the neutral and the ionised atoms.

For example, having determined $n_{0,k}$ and $n_{1,k}$ from lines of neutral and singly ionised helium (the line 4686 Å of He II), we can determine, by means of (20.18) for a given value of T, the ratio n_1/n_2, i. e. the ratio of the number of singly ionised helium atoms to the number of doubly ionised helium atoms. The greatest uncertainty here is the fixing of T. However, if the difference $\chi_{0,k} - \chi_{1,k}$ is small, an error in the value taken for T has a small effect. We can take $T = 6000°$ here also (see above). In this way, for instance, it is found that the ratio of n_2 to n_1, which is the ratio of n (He III) to n (He II), in the prominences is about $10^{38.5}$ times greater than the value given by the ionisation formula (5.11) for $T = 5040°$. From a formal standpoint, such deviations from thermodynamic equilibrium can be described as follows. In formula (5.11a), instead of the principal temperature term $5040\,\chi_r/T$, we write $\Phi\,(\chi_r)$. In this case the ionisation formula takes the form

$$\log_{10}\frac{n_{r+1}}{n_r} = \log_{10}\frac{u_{r+1}}{u_r} + \tfrac{5}{2}\log_{10} T - \Phi\,(\chi_r) - \\ - \log_{10} p_e + \log_{10}\frac{2\,(2\pi m_e)^{3/2}\,k^{5/2}}{h^3}. \tag{20.19}$$

Introducing into (20.19) the known value of p_e, taking in the second term on the right-hand side of (20.19) some definite value of T (say, close to the effective temperature of the Sun; this term is generally unimportant), and finding the ratio n_{r+1}/n_r by means of a formula similar to (20.18), we find the function $\Phi\,(\chi_r)$ from (20.19).

Using a number of elements, we can construct the relation between $\Phi(\chi_r)$ and χ_r in the form of a smooth curve. Such a curve gives, for every χ_r, a quantitative measure of the deviations from the formula of thermal ionisation (5.11)*. By means of this curve and formula (20.19), we can estimate the ionisation for elements where only n_r, or n_{r+1}, etc., is known.

7. The kinetic temperature and chemical composition of prominences. Let us now consider the question of the kinetic temperatures of prominences. One of the most direct methods of determining these is to

* See [168, p. 690] for a graph of the function $\Phi(\chi_r)$, constructed by A. Unsöld for an eruptive prominence.

study the energy distribution in the recombination emission of prominences beyond the limit of the Balmer series.

Since a prominence is transparent in the frequencies of the continuous spectrum (at least up to the limit of the Lyman series) the amount $I_\nu \, d\nu$ of energy radiated by the prominence from 1 cm³ in 1 second in a given direction cc' (see Fig. 53), and referred to unit solid angle and the frequency interval from ν to $\nu + d\nu$, must be

$$I_\nu \, d\nu = d\nu \int_0^s j_\nu \, \varrho \, ds \,, \tag{20.20}$$

where s is the geometrical length of the cylinder cc'. The quantity I_ν is evidently the intensity of radiation in the given direction.

In the case considered, the emission coefficient j_ν is given by formula (8.38), where we can neglect stimulated emission. Again supposing that the physical state of the matter in a prominence depends only on the height h, we can take the product $j_\nu \, \varrho$ in (20.20) outside the integral. We then obtain from (8.38)

$$I_\nu = s \left\{n_e \, n_1\right\} \left\{ \frac{2 \, k^2 \, h^4 \, \nu^3 \, k_\nu'}{c^2 \, (2 \, \pi \, m_e \, kT_\varepsilon)^{3/2}} \right\} e^{-(h\nu - \chi_{0,k})/kT_\varepsilon} \,. \tag{20.21}$$

For the proton $u_1 = 1$, and $g_{0,k} = 2 k^2$.

Next, taking into account the fact that $n_e \approx n_1$, and also using (5.24), we find

$$I_\nu \approx s n_e^2 \, \frac{2^6 \, \pi^2 \, h \, e^6 \, R}{3 \sqrt{3} \, c^3 \, (2 \, \pi \, m_e \, kT_\varepsilon)^{3/2}} \, \frac{g'}{k^3} \, e^{-(h\nu - \chi_{0,k})/kT_\varepsilon} \,, \tag{20.22}$$

where $k = 2$ for the Balmer continuum. Since the correction factor g' varies relatively little with frequency, it follows from (20.22) that the intensity of recombination emission beyond the series limit (for hydrogen) decreases proportionally to $e^{-h\nu/kT_\varepsilon}$. Consequently, by comparing the observations with this theoretical law, we can determine T_ε.

There is no reasonably accurate determination of T_ε by this method. The existing data on the kinetic temperatures of prominences (which are identified with the electron temperatures) are based mainly on the study of equivalent widths (using the theory of curves of growth; see above), and of the contours of emission lines in the spectra of prominences. Here we may recall that the observed width of any line is due not only to the thermal velocities, but also to the turbulent velocities of the matter in prominences. The two velocities can be distinguished if we use two elements with greatly differing atomic weights, for instance hydrogen and calcium. In this case the thermal component of the width of the line should be different, while the turbulent component is the same.

However, there is still a discrepancy between the values of T_k and v_t found by different investigators. The values found for T_k lie between 6000° and 15,000°, while the turbulent velocities are from 0 to 10 km/sec. Further investigations are needed here also. It is possible that the differences concerned may largely be the result of actual differences in the values of T_k and v_t for different prominences.

It is therefore of interest to note the following fact. It is found that the contours of prominence lines, and in particular of H_α, vary markedly from one prominence to another, and even in different parts of the same prominence, as was noticed by M. Conway and M. A. Ellison [41]; this phenomenon has recently been studied by G. S. Ivanov-Kholodnyĭ [62]. The observed broadening of the contour towards the centre of the prominence may be due to the fact that the core of the prominence has a higher kinetic temperature than its "shell".

Let us now briefly consider the problem of the chemical composition of the matter which forms prominences. There are not as yet any sufficiently detailed quantitative investigations. However, the existing general investigations of prominence spectra, and in particular those of V. A. Krat [72], show that the chemical composition of prominences is practically the same as that of the chromosphere and photosphere. Moreover, the existing data enable us to assert that, in general, all parts of the solar atmosphere (chromosphere, corona, etc.) have the same relative content of various elements. For some of these parts (e. g. the solar corona; see Chapter 22), this is confirmed directly by a quantitative analysis of spectra.

Such a homogeneity in the chemical composition of the various parts of the solar atmosphere must be a result of the "mixing" effect of *turbulent currents*. Thus, for instance, a certain difference between the spectra of quiescent prominences and those of sunspot prominences is undoubtedly a result of the differing excitation conditions, and possibly of a general difference in the number of emitting atoms.

The hypothesis of the identity of chemical composition of the photosphere and prominences reveals the possibility of determining the relative helium content in the solar atmosphere from emission lines in the spectra of prominences. As is well known, there are no helium lines in the ordinary spectrum of the Sun.

Such calculations can be performed, for example, by means of formula (20.14), bearing in mind, of course, the critical remarks pertaining to it. We write the expression for $n_{r,k}$, in accordance with (20.14), twice: once for neutral hydrogen and once for neutral helium, denoting by asterisks all quantities referring to helium. We then obtain for $n_{0,s}{}^*/n_{0,k}$

$$\frac{n_{0,s}{}^*}{n_{0,k}} = \frac{n_1{}^*}{n_1} \frac{g_{0,s}{}^*}{g_{0,k}} \frac{u_1}{u_1{}^*} e^{-(\chi_{0,k}-\chi_{0,s}{}^*)/kT} . \tag{20.23}$$

Having selected two lines with small $\chi_{0,k}$ and $\chi_{0,s}^*$, we determine, from observations of these lines, $n_{0,s}^*/n_{0,k} \approx N_{0,s}^*/N_{0,k}$, and then, introducing this ratio in (20.23), and also the most likely temperature T, we find the ratio n_1^*/n_1. Next, having estimated by some method the ionisation state of hydrogen and helium, we can estimate, from the ratio n_1^*/n_1 found, the ratio of the total number of atoms of hydrogen and helium, i. e. the ratio

$$n^*/n = (n_2^* + n_1^* + n_0^*) / (n_1 + n_0).$$

This method has been employed by various authors. A. B. SEVERNYĬ [129] used as the two lines the H_α line and the orthohelium line $\lambda = 10{,}830$ Å. The observations were carried out with the interference-polarisation filter of the Crimean Astrophysical Observatory.

As a result of such investigations it has been found (by A. B. SEVERNYĬ, A. UNSÖLD) that, in fact, the content of helium relative to hydrogen in the prominences, and consequently in the photosphere, is approximately the same as in the photospheres of stars of other classes. However, these estimates cannot really be called accurate, because of the extremely complex physical conditions in prominences.

8. A possible interpretation of the phenomena observed in prominences. Let us now briefly consider the problem of the interpretation of the chief phenomena which characterise the formation and development of prominences.

We shall discuss first the question of the mechanism whereby relatively cool and dense matter is produced, in the form of prominences, inside the very hot and rarefied corona, and also that of the converse process, the decay of a prominence. We recall that the appearance of luminous matter in the coronal region (the jets and knots in a prominence), as it were "from nowhere", is the most characteristic property of ordinary coronal prominences, coronal clouds, and a number of other kinds of sunspot prominence.

This appearance of luminous matter on the darker background of the corona is undoubtedly the result of a condensation of the matter which forms the corona, under the action of some "focusing" agencies, which appear to be electromagnetic in nature.

It is true that the kinetic temperature of the coronal matter (about 1,000,000°) is very much higher than that of the prominences (about 6000° to 15,000°). However, it is easy to explain this cooling of the coronal matter when it is condensed into prominences. For the condensation process brings about, in accordance with (8.38), an increase in the number of recombinations (increase in the product $n_e n_{r+1}$), consequently an increased loss of radiant energy into the surrounding space, and

thereby a cooling of the coronal material. In particular, it would be possible to explain, by processes of condensation of matter (of an "impulsive" short-period kind), the increases in brightness of the knots in a prominence, established by A. B. SEVERNYĬ, which coincide with sudden changes in its rate of motion.

Two possible factors must be considered in connection with the attenuation of the radiation from matter in a prominence:

(1) The expansion of the knots and of the parts of a prominence during its motion. As a result of this expansion p_e decreases, the ionisation of atoms increases and at the same time the number of recombinations (which determine, for instance, the emission in hydrogen and helium lines) is diminished. These effects lead, of course, to a reduction of the light of the prominence in emission lines.

(2) As was first established by A. B. SEVERNYĬ [128], a heating of prominences by the corona adjoining them (a conduction process) is possible. This heating also is accompanied by an increase in the ionisation of the atoms in the prominence. This factor is most effective in conjunction with the first factor, the diminution in density of the matter in the prominence. For in that case the penetration of high-velocity electrons from the corona into the prominence is considerably facilitated.

The relatively long existence of prominences of some classes is apparently brought about to some extent by the magnetic lines of force, which hinder the penetration of electrons from the corona into the prominence; these electrons have a very high kinetic temperature and can therefore ionise the prominences.

Let us now turn to the most complex problem in the physics of prominences, that of the forces which support the matter in prominences and cause it to move. Here, unfortunately, the situation is still uncertain. However, we should examine some aspects of the question.

We begin with the problem of the support of the matter which forms prominences. We might attempt to explain the relatively stable existence of some types of prominence (for example, quiescent prominences) by starting from the hypothesis that the matter in them is in a state of hydrostatic equilibrium described by equation (7.1). Writing the equation of state of the gas in the form

$$p = R\varrho\, T/\mu \tag{20.24}$$

and taking T = constant inside the prominence, as appears fairly legitimate in view of its small mean optical thickness, we obtain the equation of hydrostatic equilibrium in the form

$$\frac{d\varrho}{\varrho} = \frac{g\,\mu}{RT}\,dh\,. \tag{20.25}$$

Here we have neglected, for the present, any possible contribution due to radiation pressure.

Integrating (20.25), we obtain

$$\varrho = \varrho_0 \, e^{g\mu h/RT}, \tag{20.26}$$

where ϱ_0 is the density of matter at the base of the prominence; it must be recalled that, in our notation, h increases towards the centre of the Sun.

Taking a series of values for T which lie in a certain range, we can compare the law (20.26) with that given by observation for the density variation of any *one* element in all stages of ionisation (since all the elements are distributed in the same manner in a prominence). Such a comparison invariably shows that the extent of prominences, as regards height, is considerably greater than we should expect from (20.26). This means that some other forces directed away from the centre of the Sun, besides the gradient of the gas pressure, act against the force of gravity. It was thought for a long time that the radiation pressure, both general and selective, was such a force. An attempt was made to explain by radiation pressure not only the considerable height of prominences, but also their motion (for example, the "explosion" of eruptive prominences and their subsequent motion away from the Sun). Since this question is very important for the problem we are considering, we shall discuss it in somewhat greater detail. We shall show that radiation pressure does not seem to play an inportant part in supporting and moving the prominences.

Firstly, this follows directly from an estimate of the magnitude of the radiation pressure on the matter in prominences. Here we should be mainly interested in the radiation pressure on the atoms of the most common element, hydrogen. For, if the radiation pressure is important only for some other element, such as calcium, the atoms of calcium would have to carry (or support) an enormous number of hydrogen atoms (which are over 5×10^5 times as numerous as calcium atoms); thus the net effect of the radiation pressure on the whole of the matter would be very small.

The radiation pressure on the hydrogen atoms in prominences includes the selective radiation pressure in the frequencies of the lines of the Lyman series and the general pressure in frequencies in the Lyman continuum.

The radiation pressure on one atom of any element is easily calculated by means of the general expression (7.5). Introducing here the values

of n and k_ν from (5.20) and (5.22), we obtain for the force of radiation pressure F', acting on one atom, the expression

$$F' = \frac{\mathrm{d}p'}{n\,\mathrm{d}h} = \pi \int_0^\infty H_\nu\, k_\nu\, \mathrm{d}\nu/c\,. \tag{20.27}$$

In order that the atoms concerned should be in equilibrium under the action of radiation pressure and of the oppositely directed force of gravity, it is necessary for the equation

$$mg_\odot = \pi \int_0^\infty H_\nu\, k_\nu \mathrm{d}\nu/c\,, \tag{20.28}$$

to hold, where m is the mass of one atom of the kind in question.

Let us first consider the radiation pressure on hydrogen atoms which is determined by the Lyman continuum. Substituting $m = m_{\mathrm{H}}$ in (20.28), putting $H_\nu = B_\nu(T)$, where B_ν is given by formula (3.11), and using formula (5.24) for k_ν, we easily find that (20.28) is satisfied for $T \approx 11{,}000°$. (The lower limit of integration in (20.28) is here equal to ν_1, the frequency of the Lyman series limit.) Thus, in order to explain, in this case, the equilibrium of prominences, we should have to suppose that the Sun radiates anomalously strongly just beyond the limit of the Lyman series. [If the Sun radiated, beyond the Lyman series limit, directly from the surface of the photosphere (the chromosphere being absent), then, according to Table 13, the radiation would correspond to $T \approx 4400°$.] Such an anomaly does in fact exist (see Chapters 21 and 22), but it is much less pronounced, and corresponds to a temperature of approximately 6500°, which value is confirmed by ionospheric observations. Furthermore, at $T = 11{,}000°$ the hydrogen atoms in prominences would be mainly ionised, and this would greatly reduce the radiation pressure, since the latter acts only on neutral hydrogen atoms, and these would have to support the remaining, predominant, mass of protons.

We now pass to selective radiation pressure. Here the photospheric radiation itself is quite inadequate. In the lines of the Balmer series, the flux of photospheric radiation is known, and the radiation pressure on the atoms in the second level is easily calculated. However, the number of atoms in the second excitation state is so small that the mean acceleration acquired by one neutral hydrogen atom from the selective radiation pressure, in the frequencies of the lines of the Balmer series, is negligible compared with the acceleration due to gravity. In the region of the Lyman series, the photospheric radiation is greatly attenuated by the presence of the absorption lines in this series. A considerable

acceleration due to radiation pressure on the prominences might be given by radiation of the chromosphere in the Lyman α, β, ... lines. However, the optical thicknesses of prominences in the H_α line are usually greater than or comparable with unity. This means that the optical thickness of the prominence in the Lyman α, β, ... lines is extremely great. Thus the flux of chromospheric radiation inside the prominence is very small, and so the height of the prominence would itself be very small.

Besides these difficulties which follow from calculation, there are also a number of facts derived from observation which are in disagreement with the hypothesis that radiation pressure plays a large part in prominences. We shall briefly enumerate some of these facts.

(1) The ejection of an eruptive prominence usually takes place in regions where neither photospheric nor chromospheric disturbances are observed. If, on the other hand, this ejection took place under the influence of radiation pressure, we should expect it to occur chiefly in regions of increased intensity of ultra-violet radiation, i. e. above faculae (or flocculi) and, in general, above the active parts of the Sun (we are not speaking here of prominences connected with chromospheric flares). However, in prominences above faculae which contain sunspots, we most frequently observe *downward* motions of matter (in the majority of sunspot prominences).

(2) The formation of denser luminous matter (against the background of the darker coronal regions) in coronal sunspot prominences, in coronal clouds, etc., and its motion downwards to the chromosphere, bear no relation to the hypothesis that prominences are formed from the chromosphere under the action of radiation pressure. It is very difficult to suppose that, in the formation of a coronal cloud, the forces of gravity and of radiation pressure are in exact equilibrium*. It is clear that the mechanism whereby a coronal cloud is formed has no connection with radiation pressure.

Besides the difficulties mentioned, we might point to a number of others. Moreover, the hypothesis of radiation pressure cannot explain a large number of specific properties of the motion and behaviour of prominences. Among these properties are: (1) the motion of condensations and currents, in prominences of some classes, along, as it were, completely "fixed" trajectories, both curved and straight; (2) the existence of centres of attraction in the chromosphere (or beneath it); (3) the exchange of matter between two interacting prominences; (4) the transformation of prominences from one class to another; (5) the process of the condensation of luminous matter in coronal space; (6) the

* We recall that a coronal cloud, on its appearance, remains continually at the same height.

ejection and subsequent return of luminous matter and, what is particularly important, its returning along the same trajectory in the case of surge prominences; and so on.

Finally, we should note the following facts. The investigation of surge prominences observed above chromospheric flares (M. A. ELLISON) has shown that their motions often exhibit accelerations (towards the centre of the Sun) exceeding the acceleration due to gravity. Thus, for example, in the case of a surge prominence observed above the flare of 8 May 1951, some of the knots had an acceleration much greater than that due to gravity, an extremely important observation [17].

In general, the data at present existing show that the force of gravity plays hardly any part in the motions of many types of prominence. This seems to be true even for quiescent prominences [126].

How can we account for all the properties of prominences which we have enumerated, as well as for others? There are a number of reasons for supposing that these properties are determined in some way by the electromagnetic fields which exist in the solar atmosphere (see Chapter 18). These fields must determine the equilibrium, the motion of matter in prominences along curved trajectories, the existence of centres of attraction, the condensation processes in coronal matter, and so on.

It is true that these general considerations have as yet been little developed. However, the presence of electromagnetic fields in the Sun's atmosphere, and the large part which they play in the motions of ionised matter, are beyond the slightest doubt. The taking into account of electromagnetic fields is a quite new, and at the same time extremely important, part of the physics of the Sun.

We have as yet no well-developed theory of prominences, and hence we shall give only one of the possible models.

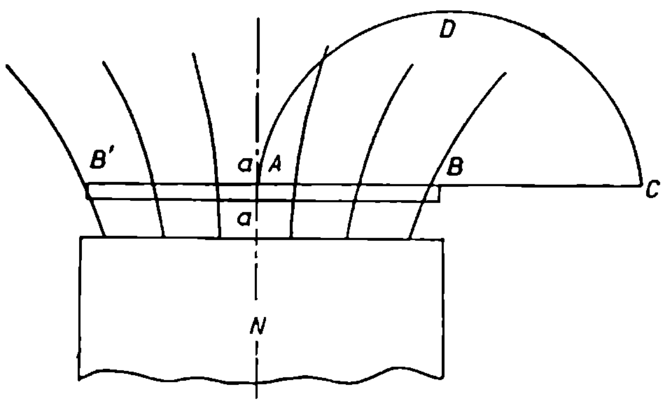

FIG. 55

H. ALFVÉN starts from what is called **unipolar induction**. This phenomenon can be explained by means of a simple model (Fig. 55). The conducting disc BB' rotates in a magnetic field about the axis aa which is parallel to the field. The motion of matter, in accordance with (18.36), produces in it a field $E' = (\mathbf{v} \wedge \mathbf{H})/c$, directed along the

radius of the disc. The potential difference between the points A and B is $V = (1/c) \int_A^B (\mathbf{v} \wedge \mathbf{H}) \,.\, d\mathbf{s}$. If we join these points by a conducting contour $ADCB$, a current will flow in it. Such a mechanism may exist on the Sun. As is well known, a motion of masses of gas is generally observed above spots (the Evershed effect), and the Coriolis force produces a tangential velocity component. This motion, which takes place under the action of hydrodynamic forces, causes a potential difference. The contour joining the points A and B is formed by a magnetic line of force ADC (since across the magnetic field the conductivity of the chromosphere is almost zero because of the low density) and the segment of photosphere CB, where the conductivity across $\mathbf{H}$ is fairly large because of the high density of matter. If we take $v = 10^4$ cm/sec, $H = 200$ oersted and $AB = 5 \times 10^9$ cm, $V = 10^8$ volts. Part of this field may cause a discharge in the photospheric layers, and the remaining part may cause a current in the chromosphere. ALFVÉN thinks that such a mechanism can explain the appearance of rapidly changing prominences whose motion ends in a spot (for example, coronal sunspot prominences). A similar process may occur every time a line of force crosses layers with different velocities in a magnetic field. For example, a potential difference may arise on account of the differential rotation of the Sun in its general magnetic field. If, as a result of the superposition of the spot field, a line of force joins two points having different velocities, a discharge may result.

The formation of prominences is pictured by ALFVÉN as follows. The current flowing along a line of force forms its own magnetic field in the form of concentric rings. It is well known that magnetic lines of force always, as it were, strive to contract. In other words, the interaction of parallel currents leads to a compression of the material forming the corona towards the axis of the current. The compression of ionised matter increases the radiation and leads to cooling. The condensations formed move along the lines of force to the photosphere under the action of gravity. The motion may also be caused by the pressure gradient formed by non-uniform compression in different parts of the line of force.

In quantitative calculations, it must be borne in mind that the density of the current flowing along the axis of the discharge is determined *not by the ohmic but by the inductive resistance*, since the self-induction of such large contours is very large.

Thus the motion of prominences seems to be a peculiar magnetohydrodynamical phenomenon. The obstacle to any complete and accurate treatment of this phenomenon lies in the lack of information concerning

the actual boundary conditions, which are necessary for a simultaneous solution of the equations of the electromagnetic field and those of the motion of the solar plasma. All attempts to discuss the motion for a given field or the field for a given motion of the highly-ionised plasma are scarcely adequate. The attempt by A. B. SEVERNYĬ [132] to consider such a complete problem in the first (linear) approximation, taking into account the force of gravitation and the gas pressure, has shown that local pressure variations (e. g. between a sunspot and its surroundings) may cause motions of plasma condensations almost exactly along lines of force in the external magnetic field of the spot if this field is large (about 1000 oersted), and at a certain angle to the lines of force for a weak field (~ 10 oersted). A. B. SEVERNYĬ has shown that, in the case of an incompressible plasma, these general solutions coincide with the well-known solutions of ALFVÉN.

It must be noticed that, although purely hydrodynamical forces cannot explain the motions of prominences, these forces cannot be neglected, since they determine the character of the motions and cause secondary phenomena. It may be that the irregular motions in quiescent prominences and in coronal clouds are determined by these hydrodynamical forces.

In conclusion, it must be emphasised that, besides the theoretical consideration of the observed phenomena which accompany the motion and equilibrium of prominences, a very great amount of observational work is required. For example, the question must be investigated to what extent the motions of luminous matter recorded in photographs are actual displacements and do not represent a simple change in position of the condensation in the corona, and consequently of the radiation. This problem must be studied by simultaneously photographing prominences and determining their radial velocities. The existing investigations seem to show that in the majority of cases we are in fact concerned with actual motions of matter, but there are as yet very few such investigations.

Furthermore, the relation between the motion of prominences and the conditions under which they radiate must be examined from every aspect, as was done in the investigations by A. B. SEVERNYĬ mentioned above.

Finally, it would be extremely interesting to ascertain whether there is any visible connection between prominences and the chromosphere (or photosphere) lying beneath them. Thus, for instance, it is known that, on H_α spectroheliograms, filaments sometimes have a bright background (in the form of a border), a few seconds of arc in width, which immediately adjoins the filament and is visible along its whole length.

Chapter 21. The chromosphere. Flocculi. Chromospheric flares

1. The observational investigation of the height distribution of emitting atoms. The chromosphere is a direct continuation of the outermost layers of the solar photosphere, which are conventionally called the reversing layer. The **base of the chromosphere** is the level in the solar atmosphere where the ordinary absorption-line spectrum of the Sun, as seen by an observer looking at the limb of the disc, passes into an emission-line spectrum like that of the prominences. Below the base of the chromosphere, the optical thickness of the solar gases along the line of sight is greater than unity in the frequencies of the continuous spectrum, and here $I_\nu(\frac{1}{2}\pi,0) \approx B_\nu(T_0)$, where T_0 is the boundary temperature of the solar photosphere. At the base of the chromosphere, this optical thickness along the line of sight is of the order of unity.

Above this lies the chromosphere itself, which is characterised by the fact that it is transparent in the frequencies of the continuous spectrum. This is its chief distinguishing property, which brings about its bright-line spectrum. Heated gases which are transparent in the frequencies of the continuous spectrum must give a *line* spectrum (if, of course, these gases are not seen in projection against a brighter background). The bright-line spectrum of the chromosphere is qualitatively similar to the spectra of prominences. In the relative intensities of a very great number of its lines, the spectrum of the chromosphere corresponds to the absorption spectrum of a star of a class earlier than that of the Sun (dG 3). In the spectrum of the chromosphere, as in those of prominences, there are lines of He I and the line 4686 Å of He II. The gases of the chromosphere are consequently in conditions which differ from those in the solar photosphere.

The study of these conditions is one of the chief tasks of the physics of the solar chromosphere. A second important task is to study the law of density variation of the matter in the chromosphere with the height x (measured from the base of the chromosphere). Here we can study this law directly by observation, which is not the case for the solar photosphere, and the same can be done for the corona. This is particularly important because, as we shall see, the distribution of the gases in the solar chromosphere is very anomalous. Thus, for example, emission in the H and K lines of Ca II can be observed up to a height $x = 15{,}000$ km, whereas, in the gravitational equilibrium described by formula (20.26) with $T = T_0 \approx 4630°$, the density of matter would have to fall practically to zero at a distance of 15,000 km, and consequently the emission of Ca II could not be detected at this height.

We shall now turn to the study of the law of density variation of the chromospheric gases with height. Let I_{ik}, or I simply, be the energy in the line corresponding to the transition $k \to i$, emitted by 1 cm² area of the chromosphere projected on the sky, in 1 second and in unit solid angle. This energy is determined by the transitions $k \to i$ taking place in the cylinder $ABCD$ (see Fig. 56a), lying along the line of sight at a height x above the base of the chromosphere and having a cross-section of 1 cm². It is evident that this energy is the total intensity in the whole line, i.e. the integral $\int I_\nu \, d\nu$ taken over all frequencies inside the line. Next, let j' be the total emission coefficient referred to 1 cm³.

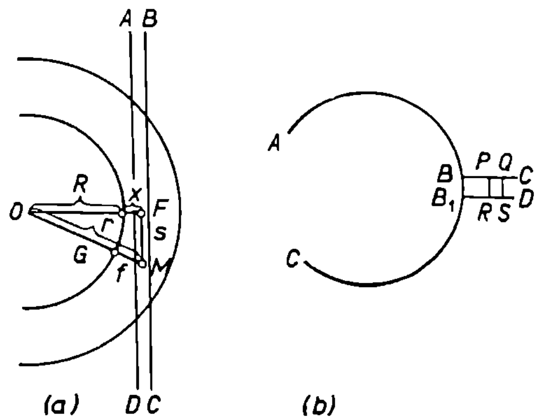

Fig. 56

The problem is to find the dependence of $j'(x)$ on x or of $j'(r)$ on r, where r is the distance of the point concerned from the centre of the Sun, using the law of variation of $I(x)$ with x which is found from observation, i. e. to convert the observed emission from unit projected area of the chromosphere into the emission per unit volume. Having found $j'(r)$, we can also find the law of variation with x or r of the number of emitting atoms.

In order to solve this problem, we make the following assumptions. We shall suppose that the chromosphere is spherically symmetric with respect to the centre of the Sun. Further, we assume that there is no self-absorption in the emission lines. (We shall discuss later how to take self-absorption into account.)

Let ABB_1C be an instantaneous position of the limb of the Moon's disc (Fig. 56b) as it gradually covers the Sun. We take 1 cm of the limb of the disc, BB_1, and denote the energy of the radiation of all atoms, projected on the semi-infinite strip C_1BB_1D, in the spectral line concerned, by $E(x)$, where x is the height of the projected limb of the Moon above the base of the chromosphere. This function $E(x)$ is found from photometric measurements of the intensity of the line at various moments during a solar eclipse, i. e. for various heights of the Moon's limb above the base of the chromosphere (the photographs of the spectrum are taken with a wide slit). It is evident that the integrated intensity $I(x)$ introduced above, which is defined in Fig. 56 (b) by the unit projected area $PQRS$, is obtained from $E(x)$ by differentiation:

$$I(x) = -\partial E(x)/\partial x \,. \tag{21.1}$$

Introducing the co-ordinate s measured along the cylinder $ABCD$ (in Fig. 56a; $s = FM$), we find for $I(x)$, in the case of a medium transparent to its own radiation, the expression

$$I(x) = \int_{-\infty}^{\infty} j'(r)\,\mathrm{d}s\,. \tag{21.2}$$

But

$$\left.\begin{aligned} s &= \sqrt{[r^2 - (R+x)^2]}\,, \\ \mathrm{d}s &= \frac{r\,\mathrm{d}r}{\sqrt{[r^2-(R+x)^2]}}\,, \end{aligned}\right\} \tag{21.3}$$

and consequently

$$I(x) = 2\int_{R+x}^{\infty} \frac{j'(r)\,r\,\mathrm{d}r}{\sqrt{[r^2-(R+x)^2]}}\,. \tag{21.4}$$

Instead of r, we introduce the height f above the base of the chromosphere ($f = GM$ in Fig. 56a). Then

$$I(x) = 2\int_{x}^{\infty} \frac{j'(f)\,(R+f)\,\mathrm{d}f}{\sqrt{[(R+f)^2-(R+x)^2]}}\,. \tag{21.5}$$

The chromosphere is of relatively narrow extent, so that j' is practically zero for $x > 15{,}000$ km. Hence throughout the chromosphere, we can neglect the quantities f and x compared with R, and thus we have

$$I(x) = \sqrt{(2R)}\int_{x}^{\infty} \frac{j'(f)\,\mathrm{d}f}{\sqrt{(f-x)}}\,. \tag{21.6}$$

This is Abel's equation; its solution is given by the formula

$$j'(x) = -\frac{1}{\pi\sqrt{(2R)}}\frac{\mathrm{d}}{\mathrm{d}x}\int_{x}^{\infty} \frac{I(f)\,\mathrm{d}f}{\sqrt{(f-x)}}\,. \tag{21.7}$$

Formulae (21.1) and (21.7) enable us to obtain the function $j'(x)$ from $E(x)$.

The reduction of observations shows that $E(x)$ can generally be represented as an exponential function

$$E(x) = A\,e^{-\beta x} \tag{21.8}$$

or as a sum

$$E(x) = A_1\,e^{-\beta_1 x} + A_2\,e^{-\beta_2 x}\,. \tag{21.9}$$

Sometimes a third term of the same form has to be added to the right-hand side of formula (21.9). Since the approximation (21.8) is very

often sufficiently accurate, we shall discuss it here. Inserting (21.8) in (21.1), we obtain

$$I(x) = A\,\beta\,e^{-\beta x}\,. \tag{21.10}$$

Finally, we introduce (21.10) in (21.7). We then find

$$j'(x) = \frac{A\,\beta^{3/2}}{\sqrt{(2\,\pi\,R)}}\,e^{-\beta x}\,, \tag{21.11}$$

and also

$$j'(0) = \frac{A\,\beta^{3/2}}{\sqrt{(2\,\pi\,R)}}\,. \tag{21.12}$$

In Table 14 we give the values of the constant β appearing in formula (21.11), and also those of $\log_{10} A$. These data were obtained by V. P. Vyazanitsyn from spectrograms of the 1941 eclipse [177].

In order to use Table 14 to draw the appropriate conclusions concerning the density distribution of elements with height, we first convert the quantity j' to the number of emitting atoms in 1 cm³. If there are n_k atoms in 1 cm³ which produce radiation in the given line with the transition $k \rightarrow i$, then, as in (20.1),

$$j' = n_k\,A_{ki}\,h\,\nu_{ik}/4\,\pi\,. \tag{21.13}$$

Table 14

Line		$\beta \times 10^8$	$\log_{10} A$	Line		$\beta \times 10^8$	$\log_{10} A$
H:	H_α	0·90	16·17	Ca I:	4227	1·33	14·08
	H_β	1·20	15·71	Ca II:	3969	0·92	15·76
	H_γ	1·30	15·48		3934	0·85	15·76
	H_δ	1·32	15·32	Ti II:	4572	2·60	13·56
	H_ε	1·30	15·15		4550	1·58	13·95
	H_8	1·42	15·07		4564	1·63	13·33
	H_9	1·37	14·75	Mn I:	4034·5		
	H_{10}	1·45	14·60		4033	1·96	13·52
	H_{11}	1·46	14·43		4031		
	H_{12}	1·47	14·45	Fe I:	4384	0·91	13·04
	H_{13}	1·52	14·44		4064	1·28	13·30
	H_{14}	1·60	14·27		4046	1·14	13·53
	H_{15}	1·55	14·15		3860	1·36	14·14
	H_{16}	1·68	14·08		3856	1·69	13·24
	H_{17}	1·70	13·75	Fe II:	5018	1·32	13·55
	H_{18}	2·04	13·60		4924	1·34	13·53
He:	5016	1·69	13·69		4584	1·23	13·30
	4922	1·32	13·49		4233	1·54	13·70
	4713	1·96	14·33		4523	1·37	12·94
	5876	0·88	14·90	Sr II:	4216	1·58	13·90
	4472	1·11	14·05		4078	1·60	14·55
	4026	0·75	13·48	Ba II:	4934	(2·05)	(14·58)
Mg:	5184	1·64	14·10				
	5173, 5167	1·22	13·98				

Consequently we have, according to (21.13) and (21.11),

$$n_k = \frac{4\pi A \beta^{3/2}}{\sqrt{(2\pi R)}\, A_{ki}\, h\, \nu_{ik}}\, e^{-\beta x}. \tag{21.14}$$

Thus, having found β and A from observation and knowing A_{ki}, we can calculate the volume concentration of atoms in the higher excitation state as a function of the height x in the chromosphere.

However, the chief interest is offered not by the behaviour of the quantity $n_k(x)$ which characterises the concentration of atoms in the kth excited state, but by that of the function $n(x)$ which characterises the concentration of all atoms (or ions) of the element concerned, in any excitation state. Since, however, the majority of the atoms (or ions) of the element concerned are usually in the ground state, it is of interest to ascertain the behaviour of the function $n_1(x)$, which gives the concentration of atoms in the ground state.

If the ratio $n_k(x)/n_1(x)$, i. e. the degree of excitation in the chromosphere, did not vary with the height x, the behaviour of the function $n_k(x)$ would be the same as that of the function $n_1(x)$, and we should know the actual height distribution of the atoms in question. It might appear at first sight that the assumption of a constant ratio $n_k(x)/n_1(x)$ in the chromosphere is justified. Thus, let us consider some transition $k \to i$, for example the transition $k \to 1$, and suppose that self-absorption is not present in this line. Next, let us assume that atoms are excited from the level i to the level k *by photospheric radiation* incident on the chromosphere; this radiation is, of course, attenuated by the corresponding absorption line caused by the transition $i \to k$. Then, because of the transparency of the chromosphere in the line considered, the density ϱ_{ik} of the exciting radiation will be practically constant throughout the chromosphere, and consequently the ratio $n_k(x)/n_1(x)$ will not change perceptibly.

However, a number of difficulties arise here, which have been discussed in detail in the previous chapter. Firstly, some emission lines in the spectrum of the chromosphere are not free from self-absorption. Secondly, the existence of lines of He I and He II in the spectrum of the chromosphere (as well as a number of other facts) indicates that we are concerned with anomalous excitation conditions, which may vary with height. In particular, it is possible that in the outer chromosphere, where the kinetic temperature is anomalously high (see the end of Section 21.3), processes of excitation by electrons play an important part, though there are reasons for supposing that the anomalies in the excitation conditions are due only to the fairly high excitation potentials (above 10 eV).

At the same time, it may be argued (see Section 21.5) that, for the majority of lines, the excitation to the corresponding levels is caused by radiation with an approximately constant density ϱ_{ik}, in which case the ratio n_k/n_1 would also be approximately constant.

The situation concerning helium and hydrogen lines is more uncertain, but here also the value of n_k/n_1 inside the chromosphere apparently cannot change by more than one order of magnitude. A more exact study of cases where $\varepsilon_k > 10$ eV requires the discovery of the actual mechanism of excitation of the atoms in the chromosphere.

2. **Self-absorption. The law of the height variation of the density of matter.** The study of the self-absorption in some lines also needs particular attention. Self-absorption distorts the law of variation of $j'(x)$, and also spuriously decreases the values of n_k. We shall therefore briefly discuss this subject.

Let us again consider the cylinder $ABCD$ (Fig. 56a), directed along the line of sight and situated at a height x above the base of the chromosphere. Let N_i be the number of atoms in this cylinder and in the lower state i of the atoms which emit by means of the transitions $k \to i$. We can then apply formula (20.9). We rewrite it as

$$\log_{10} R_{ik} = \log_{10} I_{ik} - \log_{10} P_\nu = \log_{10} \int (1 - e^{-s_\nu N_i})\, \mathrm{d}\nu \,. \quad (21.15)$$

Taking for N_i a series of continuously increasing values, we can compute the integral on the right-hand side of (21.15) for these values*. As a result we obtain a "curve of growth" relating N_i and R_{ik}. In the part of the curve where N_i is small ($X_0 = N_i s_{\nu_0} \ll 1$), the value of R_{ik} is proportional to $N_i f_{ik}$; in the intermediate part we can use an interpolation formula in which R_{ik} is proportional to $\sqrt{(\log N_i f_{ik})}$; in the last part, where N_i (and X_0) is very large, R_{ik} is proportional to $\sqrt{(N_i f_{ik})}$.

We now assume that a law of the form (21.14) holds:

$$n_i = n_i^0\, e^{-\beta f} \,, \quad (21.16)$$

where n_i^0 is the value of n_i at the base of the chromosphere, and we have introduced the quantity f in place of x in the exponent. Noticing from Fig. 56(a) that

$$f = \sqrt{[(R+x)^2 + s^2]} - R \approx x + \tfrac{1}{2} s^2/R \,, \quad (21.17)$$

we obtain for N_i

$$N_i = \int_{-\infty}^{\infty} n_i\, \mathrm{d}s = \sqrt{(2\pi R/\beta)}\, n_i^0\, e^{-\beta x} = N_i^0\, e^{-\beta x} \,, \quad (21.18)$$

* For the details of such calculations see [168, p. 626]. Here it is assumed that the absorption coefficient is determined by radiation damping and the Doppler effect. The effects of collision damping can be neglected on account of the low density of the chromosphere.

where

$$N_i^0 = \sqrt{(2\pi R/\beta)}\, n_i^0 \tag{21.19}$$

is the value of N_i at the base of the chromosphere. Thus, if formula (21.16) is valid, the law of variation of N_i (21.18) has the same exponential form.

Using (21.18) and the theoretical relation between R_{ik} and N_i, we can find the relation between R_{ik} and x, and the form of the graph of this relation is determined by the parameter β. In other words, we have a family of theoretical "curves of growth" depending on the parameter β; this parameter plays the same part as Z does in the theory of curves of growth for absorption lines (see Chapter 12).

Let us now assume that the value of $\bar{P}_\nu$ is constant throughout the chromosphere. Then, having determined from observation the value of I_{ik} for various heights x, we can construct an empirical curve of growth (connecting $\log_{10} I_{ik}$ and $\log_{10} x$), whose position as regards the axis of $\log_{10} R_{ik}$ is accurately determined apart from the term $\log_{10} \bar{P}_\nu$ (which is as yet unknown). Having obtained the empirical curve of growth, we ascertain the member of the family of theoretical curves of growth with which it most closely coincides. The parameter β is thereby determined, and hence the values of N_i^0 and $\bar{P}_\nu$. This method of determining β, N_i^0 and $\bar{P}_\nu$ is very similar to the corresponding determination of Z, N_i and v_0 in the theory of curves of growth.

The values of β and N_i^0 thus found are free from the effects of self-absorption; we can use them to determine n_i^0 from formula (21.19). This is the number of atoms in 1 cm^3 at the level corresponding to the base of the chromosphere.

The main assumption in the above method is that the excitation function $\bar{P}_\nu$ is constant throughout the chromosphere. If we knew the approximate nature of the dependence of P_ν on x, the accuracy of the method would be considerably increased. In applying this method, the change of the turbulent velocity v_t with height in the chromosphere must also be taken into account. Finally, it is necessary to bear in mind that in some cases the law (21.16) may be only approximate.

Using H_β, G. G. Cillié and D. H. Menzel [32] have found by the above method that the number of hydrogen atoms above 1 cm^2 of the base of the chromosphere, in the second excitation state, is 2×10^{12}.

However, using the same method but considerably more extensive and reliable data, obtained at the eclipses of 21 September 1941, 9 July 1945 and 25 February 1952, V. P. Vyazanitsyn (at Pulkovo) has found, from the four lines H_α to H_δ, that the number of atoms is some five times greater than the value just given, i. e. it is $\sim 10^{13}$ cm^{-2}. It is of interest that, according to his measurements, the value of β

for these hydrogen lines (taking account of self-absorption) is $1{\cdot}45 \times 10^{-8}$ cm^{-1}, whereas CILLIÉ and MENZEL give the considerably smaller value of $0{\cdot}6 \times 10^{-8}$ cm^{-1}. Such a small value of β as the latter seems very improbable.

For the H and K lines of Ca II, the value of β, allowing for self-absorption, is, according to V. P. VYAZANITSYN, $2{\cdot}1 \times 10^{-8}$ cm^{-1}, whereas MENZEL and CILLIÉ give $1{\cdot}60 \times 10^{-8}$ cm^{-1}.

The value of β found directly from observation is not usually changed by a significant amount when self-absorption is taken into consideration. The change is greatest for the strongest lines in the chromospheric spectrum, namely the H and K lines of Ca II. In this case the value of β in Table 14 is increased by a factor of approximately two. Most frequently, however, the corrections on account of self-absorption are much less. Taking these facts into consideration, as well as the remarks made above about the approximate constancy of the ratio n_k/n_1, let us consider the data in Table 14, bearing in mind formula (21.14). We compare the values of β in this Table with the values β_g which are obtained by using the formula of hydrostatic equilibrium (20.26). The latter can be rewritten

$$n(x) = n^0\, e^{-g\mu x/RT} = n^0\, e^{-\beta_g x}\,. \tag{21.20}$$

Taking as the temperature of the chromosphere the boundary temperature of the Sun $T = 4630°$, we find that for calcium $\beta_g = 284{\cdot}3 \times 10^{-8}$ cm^{-1}, and for hydrogen $\beta_g = 7{\cdot}15 \times 10^{-8}$ cm^{-1}. Thus the values of β obtained from observation (Table 14) are considerably lower than those of β_g. This is particularly noticeable on comparing the values of β and β_g for calcium, where $\beta_g \approx 190\,\beta$ (when self-absorption is taken into account, β for calcium is about $1{\cdot}50 \times 10^{-8}$).

This fact cannot be explained as the effect of ionisation increasing with height. It can be deduced from the intensities of the resonance lines that the proportion of neutral calcium atoms to ions in the chromosphere is negligibly small, and hence an increase in the number of ions at the expense of the number of neutral atoms cannot have any marked effect on the ion density. A decrease in the number of Ca II atoms with height, because of ionisation, can only increase the observed value of β. The same applies to hydrogen. It is true that there is reason to suppose that the anomalous distribution of hydrogen (small β) is a result of the anomalous radiation conditions, and in particular of the increase in the number of recombinations with height, owing to the increasing ionisation of hydrogen. However, the chemical composition of all the solar envelopes and objects in them must be the same, in consequence of the existence in them of turbulent mixing currents (see Chapter 20). In that case, however, the value of β for hydrogen, taking account of all H atoms

(neutral and ionised) should be the same as for the other elements, and in particular Ca. But here the observed value of β for H is considerably less than β_g, which is $7{\cdot}15 \times 10^{-8}$ cm^{-1}. Consequently, some additional forces, besides the ordinary elastic forces in a heated gas, must act in the chromosphere against the force of gravity. However, before considering a possible interpretation of this fact, we shall discuss some other data which characterise the physical state of the chromospheric gases.

3. **The electron concentration. The electron temperature.** Let us first consider the subject of the electron concentration n_e. Firstly, the value of n_e may be determined from formula (15.1). We shall discuss a little later the value of T to be used in applying (15.1) for the various layers of the chromosphere. Secondly, we can determine n_e from formula (20.16), assuming, as there is reason to do (see the next section), that the majority of the free electrons in the chromosphere are due to the ionisation of hydrogen. Thirdly, we can also determine n_e from formula (20.22). It follows from this formula that the emission coefficient j_ν' for recombination radiation, referred to 1 cm^3, is

$$j_\nu' = n_e^2 \frac{2^6 \pi^2 h e^6 R}{3\sqrt{3}\, c^3 (2\pi m_e k T_\varepsilon)^{3/2}} \frac{g'}{k^3} e^{-(h\nu - \chi_{0,k})/kT_\varepsilon} . \tag{21.21}$$

Having determined the value of $j_\nu'(x)$ for the Balmer (or Paschen) continuum from observation by means of formula (21.11), we can find n_e from formula (21.21), if we know T_ε. (By this method we can also determine n_e for prominences, if we know the value of s in formula (20.22), which is the thickness of the prominence in the direction considered.) It is best to do this for frequencies close to the edge of the band, where $h\nu \approx \chi_{0,k}$, and consequently the error in the value taken for T_ε appears only in the denominator of the right-hand side of (21.21), in the quantity $T_\varepsilon^{3/2}$. Here also it is assumed that the number of free electrons in the chromosphere is determined mainly by the ionisation of hydrogen. If the electrons formed by the ionisation of metals are taken into account, the value of n_e is increased.

An estimate of the electron concentration is also possible by H. Zanstra's method, which we have explained in connection with prominences. If the continuous radiation of the chromosphere in the visible region of the spectrum is due to electron scattering, while the radiation just beyond the limit of the Balmer series is due to recombinations, then the ratio of the emission coefficients found for the two processes from observation gives n_e if T_ε is known. However, the problem of which heights in the chromosphere are those where the scattering of solar radiation in the visible part of the spectrum by electrons is the chief factor that determines the observed intensity is not yet solved. To resolve this question, careful measurements of polarisation are needed.

It has been found by the above methods that n_e at the base of the chromosphere is about 2×10^{11} to 5×10^{11}, and for $T_\varepsilon \approx 5000°$ the value of p_e is about 0·1 to 0·3 bar. As the height x increases, the electron concentration n_e decreases comparatively slowly. At a height x=4000 km, the value of n_e is of the order of 2×10^{10}; at a height x = 10,000 km it is of the order of 2×10^9, and at a height x = 13,000 km it is of the order of 10^9. These estimates are some "average" values. According to J. H. PIDDINGTON (see below), the values of n_e in the years of minimum solar activity are less than the mean values given (for the same heights, of course).

Let us now turn to the next parameter, the electron temperature T_ε. One of the possible methods of determing T_ε for the chromosphere, as for the prominences, is to study the energy distribution in the recombination radiation of hydrogen beyond the limits of the Balmer and Paschen series. The starting-point is formula (21.21), according to which $j_\nu' = C\exp(-h\nu/kT_\varepsilon)$, where C is some constant. By comparing this law with the observations we find T_ε. Measurements carried out by D. H. MENZEL and G. G. CILLIÉ according to this method give a value of about 5000° for the base of the chromosphere. V. A. KRAT [72], using the same method, has obtained $T_\varepsilon \approx 6200°$ for heights of 7000 to 8000 km. In a recent paper [12] R. G. ATHAY, D. E. BILLINGS, J. W. EVANS and W. O. ROBERTS, using the same method, give for the base of the chromosphere a value of T_ε which is close to 4000°.

Other data also point to a relatively low T_ε at the base of the chromosphere. For instance, a high T_ε in the lower layers of the chromosphere would lead to a very great excess of ultra-violet chromospheric radiation, and this contradicts ionosphere observations. Moreover, if there were a high T_ε at the base of the chromosphere, we should certainly observe some forbidden lines in the spectrum of that region; these in fact do not exist. A value of 30,000° for T_ε at the base of the chromosphere, which has been found by R. O. REDMAN and confirmed by many authors, would lead to the result that the spectrum of the chromosphere would resemble that of a B or O-type star, and this is completely contrary to observation.

A second method, which can give a definite idea of the variation of T_ε within the chromosphere, is the study of the radio emission of the Sun. The general basis of this method will be considered in the next chapter. Here we shall briefly explain the results of the observations concerned.

Unfortunately, it is impossible to determine T_ε for the lowest layers of the chromosphere in the radio wavelengths hitherto used (because of the large absorption of radio waves). Fairly correct results are obtained for heights above 3000 to 5000 km. According to the recent

work of J. H. PIDDINGTON, who takes account of all existing radio observations [123], the kinetic temperature of the chromosphere at a height of 5000 km is about 6000°, i. e. the value of T_ε hardly changes over a distance of 5000 km from the base of the chromosphere. At greater heights, the value of T_ε begins to increase. At a height $h = 10{,}000$ km it is about 30,000°, while at a height $h = 14{,}000$ km we should expect the temperature to exceed 200,000°.

The variation of the temperature T_ε with height might be studied from theoretical considerations, some hypothesis being made concerning the heating of the chromosphere (e. g. thermal conduction from the corona), or concerning the means by which the solar chromosphere is supported. However, a consideration of the existing theories shows [123] that they involve many arbitrary assumptions and contradict the results of radio and other observations.

We have mentioned above that, in the range from $h = 0$ to $h \approx 5000$ km, the temperature T_ε varies very little. On the other hand, ATHAY, BILLINGS, EVANS and ROBERTS (see above) find that, even between $h = 0$ and $h \approx 2400$ km, there is a fairly rapid increase in T_ε. Their measurements of the intensity in the Balmer continuum comprise too small a range of wavelengths (λ 3500 to 3647 Å), however, and are therefore liable to considerable errors. Thus the very important problem of the gradient of the temperature T_ε in the lower layers of the solar chromosphere needs further careful study.

4. **The ionisation of atoms.** Let us now consider the subject of the ionisation of the atoms in the chromosphere. It is quite evident that, because of the relatively low value of n_e in the chromosphere, the degree of ionisation in it must be higher than in the outer layers of the photosphere, by (5.11). In reality it is even higher than the value found by applying (5.11), as may be seen merely from the presence, in the spectrum of the chromosphere, of the line 4686 Å of He II; and the observed anomalies are the greater, the greater the corresponding ionisation potential. We recall that the same is true of the prominences.

Let us consider the problem of the ionisation of hydrogen in the lower layers of the chromosphere (at heights up to 2000 km). We shall first show that the electrons in the chromosphere are in fact formed by the ionisation of hydrogen. From observation, using the considerations given above regarding self-absorption, we can calculate the concentration of Ca II atoms at a height $x = 0$ which are in the first (ground) state, i. e. the number of such atoms in 1 cm^3. According to MENZEL and CILLIÉ, as well as to a recent investigation by V. P. VYAZANITSYN (see above), it is approximately 3×10^8. Moreover, this number can as usual be approximately identified with the concentration of Ca II atoms in all states of excitation.

On the other hand, a comparison of the spectrum of the lower layers of the chromosphere with the spectra of ordinary stars shows that the former is similar, roughly speaking, to the spectra of stars of the class gF 5, as regards the relative intensity of lines of metals. (It is not possible to set up an exact correspondence here, since, for example, the class of these layers as determined from lines of metals is later than gF 5, but as determined from helium lines it is earlier.) In the atmospheres of stars of this class, the metals are chiefly singly ionised. Thus there is every reason to expect that, in the lower layers of the chromosphere, the metals are predominantly in the singly ionised state. Hence it follows, in particular, that the number of free electrons in 1 cm^3, formed by the ionisation of calcium, is equal at the base of the chromosphere to the value just mentioned, 3×10^8. Using this figure, and supposing that the atoms of the metals are predominantly singly ionised, we can calculate, by means of Table 1, the number of free electrons formed by the ionisation of all metal atoms. [The most convenient element for such calculations is iron, whose second ionisation potential is relatively very high (16·5 eV). The assumption that $n(\text{Fe III})/n(\text{Fe II}) \ll 1$ is even more justified than for calcium, but there are as yet no reliable values for $n(\text{Fe II})$.] The numbers thus obtained are smaller, by one or two orders of magnitude, than the value of n_e found directly from observation. Furthermore, the above value $n_1(\text{Ca II}) = 3 \times 10^8$ is exaggerated, since, in finding it, the increase of the turbulent velocity with height in the chromosphere was not taken into account. Consequently, there must be some additional source of electrons. It is clear that this can only be hydrogen. Thus we can suppose that $n_e \approx n(\text{H II})$, where $n(\text{H II})$ is the number of protons in 1 cm^3.

Starting from these considerations, and from the constancy of the relative chemical composition in all layers of the solar envelopes, we can calculate the degree of ionisation of hydrogen in the lower chromosphere. Taking $n(\text{Ca II}) = 3 \times 10^8$, in accordance with the above, as almost the total number of calcium atoms in 1 cm^3 at the base of the chromosphere, we can calculate $n(\text{H})$, the total number of hydrogen atoms in 1 cm^3, by Table 1. On the other hand, we know $n(\text{H II}) \approx n_e$. Consequently, we can immediately estimate the ratio $n(\text{H II})/n(\text{H})$, i. e. the degree of ionisation x_{H} of hydrogen at the base of the chromosphere. For other heights (in the lower layers of the chromosphere) we can determine $n(\text{Ca II})$ from formula (21.16), where β must be calculated with allowance for self-absorption. The corresponding calculations show that at the base of the chromosphere the value of x_{H} is approximately 0·02 to 0·03, while at a height $x = 1500$ km it is about 0·5. At still greater heights, the value of x_{H} is close to unity. Thus, in the lowest layers of the chromosphere, hydrogen is mainly neutral, yet

by reason of its high percentage content it provides almost all the free electrons. However, even here its ionisation is a hundred times greater than that estimated by formula (5.11) with $T = 5000°$ and $n_e \approx 2 \times 10^{11}$ cm^{-3}. As the height increases, so does this anomaly. It is clear that with increasing height the equality $n_e \approx n$(H II) becomes more and more exact, so that throughout the chromosphere the electrons are formed mainly by the ionisation of hydrogen.

This method is based on the supposition that the chemical composition of the whole of the solar envelopes is constant. For metals, the similarity of the chemical composition of the photosphere and of the chromosphere is directly confirmed by observation (D. H. Menzel, V. P. Vyazanitsyn and others). In particular, the presence in the chromosphere of heavy atoms such as La, Ce, Pr, etc., shows that mixing currents play a very great part in the chromosphere.

Let us now consider helium. The total number of helium atoms in 1 cm^3 in the lower layers of the chromosphere can be found by the same method, using the known values of n(CaII) $\approx n$(Ca) and Table 1. Here it is again assumed that the chemical composition of the chromosphere is the same at all heights and is identical with that of the photosphere.

The concentration n(HeII) can be found by means of formula (20.14). We can consider orthohelium as an element with an ionisation potential of 4·7 eV, since the triplet and singlet systems of helium do not combine. For such a comparatively low value of the ionisation potential (and therefore of the excitation potentials), the application of the "equilibrium" formula (20.14) ought not to involve noticeable errors, since the anomalies begin only for energies of the order of 10 eV. Thus, having determined the values of $n_{0,k}$ from the absolute intensities of the helium triplets, we can determine $n_1 = n$(HeII) in accordance with (20.14), taking the values given above for n_e and putting T equal to about 5000° or 6000°. Finally, we can obtain the ratio n(HeIII)/n(HeII), in the same way as for the prominences, from formula (20.18) where $n_1/n_2 = n$(HeII)/n(HeIII). Here, of course, we must recall the critical remarks that were made concerning the applicability of formula (20.14).

On the basis of the above considerations, we can calculate the ionisation of helium at various levels in the chromosphere. The corresponding calculations, using the observational data, lead to results similar to those obtained for hydrogen:

(1) Up to heights of some thousands of kilometres, the ratio n(He)/n(HeII) is several orders of magnitude, i. e. at these heights helium is mainly neutral.

(2) The degree of ionisation found for each layer of the chromosphere is higher than follows from the ionisation formula (5.11), and

the excess is greater for the second ionisation than for the first (the anomaly increases with χ_r).

(3) The increase of the ionisation with height takes place more rapidly than would be brought about by the decrease of n_e with height.

Thus, both for hydrogen and for helium, we have to explain, firstly, the anomalously high ionisation of the atoms of these elements and, secondly, the increase in this anomaly with height in the chromosphere (the same anomalies must hold for ionised metals whose second ionisation potentials are fairly large). The facts which we have enumerated cannot be explained by means of an anomalously high kinetic temperature in the chromosphere which increases with height, since, as we have seen, the kinetic temperature of the lower layers of the chromosphere is relatively low. Collisions can be important only in the outermost layers of the chromosphere.

According to I. S. Shklovskiĭ, these facts can be explained to a considerable extent (as they can for prominences) by the ionising power of the ultra-violet radiation of the solar corona. The hydrogen in the outer layers of the chromosphere will be ionised by the hard monochromatic radiation in the coronal lines (Ne VIII, λ 776 Å; Mg X, λ 625 Å). For helium, both the values of the ratios $n(\mathrm{He})/n(\mathrm{He\,II})$ and $n(\mathrm{He\,II})/n(\mathrm{He\,III})$ at the base of the chromosphere and the law of their variation with height are explained by photo-ionisation by the "continuous" component of the ultra-violet radiation of the corona, which is formed by the recombination of free electrons with coronal protons and α particles, and also by the bright lines in the far ultra-violet region of the spectrum of the corona.

However, preliminary calculations have shown that the ionisation by coronal radiation is apparently not great enough to explain the high value of the ratio $n(\mathrm{H\,II})/n(\mathrm{H\,I})$ for hydrogen in the lower chromosphere. This is because the optical thickness of the lower chromosphere beyond the limit of the Lyman series (where the probability of photo-ionisation of hydrogen is greatest) is very great, and the corresponding coronal radiation passes through these layers in a much attenuated form.

Later and more accurate calculations show that this difficulty may not be insuperable. In this connection, the following possibility should be pointed out. Observations show that the chromosphere is not a continuous structure. It consists of separate rising "stalks", which, in particular, have caused observers to compare the external appearance of the chromosphere (viewed "sideways") to that of a burning prairie. (This similarity is emphasised by the circumstance that these "stalks" are not steady features. They frequently change their inclination to the normal, and in general the fine structure of the chromosphere, like granulation, is continually changing.) On the other hand, observa-

tions carried out in coronal lines (chiefly the line 5303 Å) show that heated matter is ejected from the chromosphere into the corona, and this matter has the temperature of the corona, i. e. one of the order of a million degrees. It is quite possible that this matter also fills part of the space between the chromospheric "stalks" mentioned above. In view of its small density, this "coronal" matter cannot appreciably alter the general spectrum of the chromosphere. However, electrons having the very high mean kinetic energy which corresponds to a temperature of a million degrees may penetrate into the chromospheric stalks (at least their outer parts) and there cause the anomalous ionisation of hydrogen and helium. This hypothesis, however, requires further development.

5. **The mechanism of the excitation of atoms.** Let us now consider what processes determine the radiation of the chromosphere in the individual spectral lines. The difficulty of solving this problem is redoubled by the fact that we do not know the agency which brings about the hydrogen ionisation in the lower layers of the chromosphere. This agency must have an effect on the excitation processes for other atoms also. The complexity of the problem is shown by the fact that the intensities of the helium lines during some eclipses have maxima at a definite height above the chromosphere, decreasing above and below this height.

There are three conceivable sources of the chromospheric radiation: (1) the scattering of photospheric radiation by the atoms in the chromosphere; (2) recombination processes with subsequent cascade transitions; (3) the excitation of atoms by collisions.

We may suppose that, at least in some cases, and particularly for the higher levels of hydrogen and for helium, the chief source of the radiation of atoms in the lower layers of the chromosphere is recombination processes. This is confirmed by the fairly intense radiation of chromospheric hydrogen beyond the limit of the Balmer series; this radiation, of course, is due entirely to recombination. Moreover, the scattering of photospheric radiation can only give considerably less energy, for helium, than the amount observed in the corresponding lines, while collisions in the lower and middle chromosphere, in consequence of the relatively small T_ε (and n_e), cannot be important.

One of the principal means of solving this problem is the study of the distribution of the emitting atoms among quantum states. For hydrogen, in particular, this is a question of the relative intensity of the lines in the Balmer (and Paschen) series. Having found from observation the sequence of ratios $I(H_\alpha) : I(H_\beta) : I(H_\gamma) : \ldots$, we have to compare this sequence (the **decrement**, as it is called) with that given by various theories. The latter must be based on the solution of equations

of the type (20.11) to (20.13). However, there are serious difficulties here also. Firstly, for many transitions we do not know the values of ϱ_{ik} and C_{kf} which appear in these equations. Secondly, the first few members of the Balmer series (H_α, H_β, H_γ) are distorted by self-absorption, for which no exact allowance has yet been made. (The self-absorption in the H_α and H_β lines has the result that their maximum intensity varies very slightly with the height between $x = 0$ and $x = 3000$ to 4000 km.) Besides these, there are a number of other complicating circumstances, particularly the important part that collisions may play in excitation processes in the outer chromosphere (high T_ε), and so on.

A spectrophotometric investigation of the three eclipses of 21 September 1941, 9 July 1945 and 25 February 1952 has enabled V. P. Vyazanitsyn to find the observed distribution of hydrogen atoms among the excitation states with $k \geq 2$. It was found that this distribution can be represented by the Boltzmann formula (5.54) with $T_{ex} \approx 5500°$. However, the same observations indicate that for the higher levels there is, apparently, an additional excitation, corresponding to a temperature higher than 5500°. It is possible that this corresponds to recombination processes.

As regards the excitation of levels with fairly low values of the quantum number, the above value of 5500° is somewhat too large for a mechanism of simple re-emission. We have to bear in mind that the scattering of solar radiation by chromospheric atoms occurs in this case in the frequencies of absorption lines, where the intensity of this radiation is reduced by the presence of the Fraunhofer absorption lines H_α, H_β, ... On the other hand, the temperature T_E of the radiation which corresponds to the central parts of these lines is low, of the order of 4000°.

For this reason, the problem of the mechanism of the emission in the chromospheric lines H_α, H_β, ... is best resolved by a study of the *contours* of these lines, using a high dispersion, which can very well be done outside eclipse.

The contours of the first few lines of the Balmer series in the spectrum of the chromosphere have a characteristic form: in the centre of each emission line we find a fairly wide *depression*, bounded on each side by "peaks". The width of this depression is approximately equal to the width of the dark core of the corresponding Fraunhofer line H_α, H_β, ... This immediately suggests that in the present case we are concerned with the *scattering* of the solar radiation by chromospheric hydrogen [74]. This conclusion has been confirmed by V. A. Krat and T. V. Krat [75] by direct spectrophotometry of the chromospheric lines H_α and H_γ to H_9. It has been shown in the latter paper that the contours of all these lines are well explained by the operation of a re-emission mechanism.

The supposition that the radiation in the first few lines of the Balmer series in the spectrum of chromospheric hydrogen is due to processes of scattering is a wholly natural one. However, it must be borne in mind that the efficiency of this process of scattering depends on the number of hydrogen atoms in the chromosphere that are in the second excitation state. For the process to be effective, it is necessary that the number N_2 should be fairly large, i. e. the density ϱ_{12} of the chromospheric radiation in the Lyman α line should be fairly high. It is possible that recombinations may be very important here. Their number may be insufficient to produce the radiation in the lines H_α, H_β themselves, yet at the same time the density ϱ_{12} of the radiation in the Lyman α line may be high enough (on account of the very great optical thickness of the chromosphere in this line).

The following remarks should be made concerning the above. Firstly, we must bear in mind that these conclusions are valid only for the lower layers of the chromosphere. In the highest layers, the radiation, even in the first few lines of the Balmer series, may be largely determined by processes of recombination and of the collision of atoms with electrons.

Secondly, from a formal point of view we could examine the contours of chromospheric lines by using formula (20.3), where $j_\nu/\sigma_\nu = P_\nu$. If now the value of P_ν in this formula decreases with increasing height in the chromosphere, and does so fairly rapidly, it is easy to show (see Section 21.7) that the contour of an emission line will have in the centre a depression similar to that which we observe in the first few chromospheric lines of the Balmer series.

The existing observations show that, in the central parts of the Fraunhofer lines H_α, H_β, ..., where the absorption coefficient is fairly large, the intensity of the radiation diminishes as we go from the centre to the limb of the Sun's disc. This, in accordance with formula (4.27), where we must replace B_ν by P_ν, shows that in the chromosphere P_ν does in fact decrease as the height increases. However, in order to elucidate the part played by this factor in interpreting the observed contours of the chromospheric lines H_α, H_β, ..., a special investigation would be required.

For the metals, the relative part played by recombination radiation and the scattering of photospheric radiation is not yet completely clear. The latter factor is apparently the more important one. Thus, if the excitation of atoms is effected by photospheric radiation, the ratio of the equivalent widths of the same lines in the "reversing layer" and in the chromosphere (emission lines for the chromosphere) should not depend on the excitation potential of the lower levels of the lines. Measurements made by V. A. Krat [72] for lines of Fe I and Ti II (from material obtained at the 1945 eclipse) have shown that there is in fact

no dependence. However, a great amount of further work is of course needed here.

6. **The equilibrium of the chromosphere.** We must now consider the equilibrium of the chromosphere. We have seen that some additional forces, directed away from the centre of the Sun, i. e. outwards, act in the solar chromosphere, besides the ordinary forces of the elasticity of a heated gas and of gravitation.

The first question which arises here is whether the action of these forces extends to the entire chromosphere (and possibly to the outer layers of the photosphere) or is localised in some part of it. As I. S. SHKLOVSKIĬ [148] has pointed out, there is reason to suppose that, in the lowest layers of the chromosphere, up to heights of the order of 1000 km, the additional forces play practically no part.

The values of β given in Table 14 are some averages with respect to height. At the same time, as we have already said, observers very often have to introduce two terms in the formula for $E(x)$, so that $E(x)$ takes the form (21.9). The form of the dependence of $n(x)$ on x changes correspondingly. From the data of a number of observers [168, p. 620], the corresponding formula for $n(x)$ can be written (as an average for many elements)

$$n(x) = n^0(0{\cdot}982\, e^{-5{\cdot}9 \times 10^{-8} x} + 0{\cdot}018\, e^{-2{\cdot}6 \times 10^{-8} x}) \,. \tag{21.22}$$

R. WILDT gives a similar formula for Fe and Ti, obtained from a number of eclipses. On the average (taking into account the errors of observation, etc.) it is found that the value of the exponent β_1 in the first term of expressions like (21.22) is close to β_g for hydrogen with $T \approx 4600°$ to 5000°. This can be interpreted as showing that the hydrogen in the lowest layers of the solar chromosphere is supported by the ordinary forces of elasticity of a heated gas. The metals follow the same law, in consequence of the mixing of the solar gases. It must be noted that the values of β found from observation for He and H in the lowest layers of the chromosphere may differ very markedly from the β for metals. In fact, the emission of both hydrogen and helium in the chromosphere is completely anomalous, so that a relatively slow decrease of the emission in the lines of these elements with height can be wholly explained by the importance of recombinations increasing with height. This hypothesis is supported by the presence of a maximum in the helium emission at a certain height above the base of the chromosphere. It is clear that this maximum is the result of two oppositely acting factors: the increase in the emissive power of the helium atoms with height and the decrease in the concentration of helium atoms with increasing height.

If the above considerations regarding the equilibrium of the lowest layers of the chromosphere are confirmed by further investigations, we can state with complete certainty that the height distribution of the photospheric gases also is determined mainly by the ordinary elasticity of a heated gas and by gravity [equation (17.1)].

All that has been said above refers to heights of up to 1000 km. Beginning from heights of the order of 1000 km, some additional forces come into play supporting the chromosphere; here we must again emphasise that the support of the chromosphere cannot be explained by substituting a sufficiently high temperature T in the law of hydrostatic equilibrium (21.20). We have seen that the kinetic temperature of the lower, and even of the middle, chromosphere is in fact relatively low.

Let us discuss the nature of these forces. E. A. MILNE suggested that the radiation pressure on the Ca II atoms was the chief force supporting the chromosphere. Calculations carried out on the basis of formula (20.28) show that the radiation pressure on the Ca II atoms in the outer layers of the chromosphere almost exactly balances the force of gravity on these atoms. However, this theory has proved quite inadequate, for a number of reasons, the chief of which is that the number of calcium atoms in the chromosphere (and everywhere in the solar envelopes) is less than the number of hydrogen atoms by a factor of 5×10^5.

Many authors relate the forces supporting the chromosphere to turbulent processes in it. Investigations of the widths of chromospheric lines, performed by many observers (O. A. MEL'NIKOV, V. P. VYAZANITSYN, A. UNSÖLD and others), show that these widths are usually greater than would be expected if only the thermal motions of atoms were present. From this, the conclusion has been drawn by W. H. MCCREA that the matter in the chromosphere is in turbulent motion, its mean velocity being of the order of 10 to 20 km/sec. The complex filamentary structure of the outer layers of the chromosphere also indicates the presence of turbulent motions in them. There is reason to suppose that the turbulent velocity increases with the height x. This is shown by the fact that faint lines, in which the emission is observed up to heights $x \approx 1000$ km, show turbulent velocities of not more than 2 km/sec.

Let us now turn to formula (21.20). The greater the quantity RT/μ, i. e. the greater the mean velocity of the particles, the greater will be the extent of the atmosphere for a given counteraction of the force of gravity. If a turbulent motion, with a velocity distribution $e^{-(v/v_t)^2}$ (for one component), is added to the thermal motion of the particles, then, according to (11.29), equation (21.20) becomes

$$n(x) = n^0 \, e^{-gx/[(RT/\mu) + \frac{1}{2}v_t^2]} \,. \tag{21.23}$$

In this case the quantity β in the exponential law (21.16) takes the form

$$\beta = g\left\{\frac{RT}{\mu} + \frac{v_t^2}{2}\right\}^{-1}. \qquad (21.24)$$

Putting $\mu = 1$ on account of the large hydrogen content, and also $T = 4800°$, $v_t = 15$ km/sec, we obtain $\beta = 1{\cdot}8 \times 10^{-8}\,\mathrm{cm}^{-1}$, which, generally speaking, is close to the values of β in Table 14, given by V. P. Vyazanitsyn. (For hydrogen and helium the values of β, as we have said, are distorted by the dependence of the emissive power on the height. Moreover, strong lines, such as the H and K lines of Ca II, are distorted by self-absorption.)

This mechanism of support of the chromospheric gases has been discussed by a number of authors; A. B. Severnyĭ [127] considers a turbulent velocity v_t which varies proportionally to the height x. Instead of using equation (7.1), he starts from the following equation of equilibrium:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\varrho\, v_t^2 + R\,\varrho\, T/\mu\right) = -g_e\,\varrho\,, \qquad (21.25)$$

where $\varrho\, v_t^2$ are the Reynolds stresses and g_e the effective acceleration due to gravity (in the sense that it takes into account possible effects of radiation pressure). The integration of (21.25) gives the relation between the density ϱ and the height x.

The chief problem as regards this mechanism concerns the nature of the forces which maintain a steady turbulent motion in the chromosphere, i. e. it is necessary to know whence the energy comes which goes to maintain the turbulent motions. In their work on the nature of the forces which cause turbulence in the chromosphere, A. I. Lebedinskiĭ and L. É. Gurevich [55], after mentioning the inadequacy of the sources of turbulent motion previously put forward, propose a new mechanism, in which the principal forces are those of magnetic pressure (see Chapter 18). These forces arise on account of the motion of granules in the general magnetic field of the Sun. Here it is found that the extent of the chromosphere depends on the intensity of the Sun's general magnetic field. However, judging by existing data, the magnetic field varies noticeably with time, and this conflicts with the relative constancy of the height of the chromosphere.

Without rejecting entirely the mechanism of turbulent support of the chromosphere, we must also notice another possibility. We have said that the chromosphere consists of separate jet-stalks, which are similar to small prominences in their external appearance. On the other hand, we have seen that the forces which support the prominences are, according to all existing data, purely electromagnetic in character;

turbulent motions play practically no part in supporting prominences (at least of some types). This may be seen from the numerous sub-classes of prominences which arise in coronal space and move downwards.

If there is in fact a physical kinship between the prominences and the chromospheric jets (this is confirmed, for instance, by the chromospheric spicules which have recently been studied by E. YA. BUGOSLAVSKAYA, W. O. ROBERTS and others), then the chromospheric matter too may be supported by the same forces as support the prominences.

It seems quite evident from what has been said above that the study of the fine structure of the chromosphere is very important. Experience at the Pulkovo observatory has shown that a great deal of information on this subject can be obtained by studying the chromosphere outside eclipse, using a narrow-band interference-polarisation filter.

The corresponding observations, carried out by V. A. KRAT and I. A. PROKOF'EVA [76] in the H_α line with a transmission width of 0·6 Å, have shown that the chromosphere is not homogeneous. The separate jets in it form a "chromospheric network", whose cells have dimensions (horizontally) of from 8000 km to 18,000 km. The individual jets are about 2000 km thick.

Similar work by I. A. PROKOF'EVA has shown that, in the different chromospheric jets, the density gradients (i. e. the parameters β giving the decrease of density with increasing height) are very different; this makes the chromospheric jets still more akin to prominences.

All these results show that the gradient found from chromospheric observations partly reflects the *statistics* of the height distribution of the various jets, and not only the density distribution of matter in the chromosphere. This fact should be noted.

It is also very interesting that the gradients β of intensity of chromospheric lines above *facular plages* [69] are little more than half their values in the undisturbed chromosphere. The chromosphere is higher above facular plages. This means that the structure of the chromosphere above active regions differs from its structure above ordinary undisturbed parts of the Sun. The prominence-jets in the chromosphere above faculae are larger, and their mean altitude greater, than for other regions of the Sun.

So far, all our conclusions about the chromosphere have been based on observations made at the limb of the solar disc. Let us now consider chromospheric phenomena observed in projection on the solar disc.

7. The physics of flocculi. If we adjust the sufficiently narrow secondary slit of a spectroheliograph on the centre of some strong line in the solar spectrum, we can study on the spectroheliograms the chromosphere projected on the solar disc, since the central parts of some very

intense absorption lines in the solar spectrum are formed by the chromospheric layers. Such observations show that the darkening of the disc to the limb is less in the chromospheric radiation than in the parts of the continuous spectrum adjoining the lines in question.

However, the most interesting study is that of the chromospheric radiation in active regions of the Sun's surface. We shall consider here only some problems of a theoretical character, referring the reader to the appropriate sources [70; 179, p. 233; 168, p. 704] for the descriptive side of the subject.

The H and K lines of Ca II and the H_α line are those most often used in studying active regions in chromospheric radiation, and we shall now discuss these.

Spectroheliograms taken in the central parts of these lines show bright areas in active regions of the Sun. (Separate bright luminous points may be observed outside active regions also. We shall not discuss these luminous points, however.) This is easily understood if we consider the contours of the H, K and H_α lines in the active regions of the Sun, which mainly coincide, on the solar surface, with faculae (including sunspots, of course).

The contour of the H_α line at points where there are faculae is shown in Fig. 51 by a continuous line, while the chromospheric radiation proper should, according to what was said at the end of Chapter 19, be determined by the area abc. The contour bc (dotted) is the hypothetical contour of the H_α line in the spectra of faculae.

The intensity at the point c in Fig. 51 is somewhat indeterminate, since we do not know the magnitude of ac, i. e. the energy due to the chromosphere proper. The same indeterminacy holds for the undisturbed parts of the photosphere also. Here, too, we do not know the relative importance of the chromosphere and of the outermost parts of the "reversing layer" in producing the observed central intensities of the H_α, H_β, H_γ, ... lines.

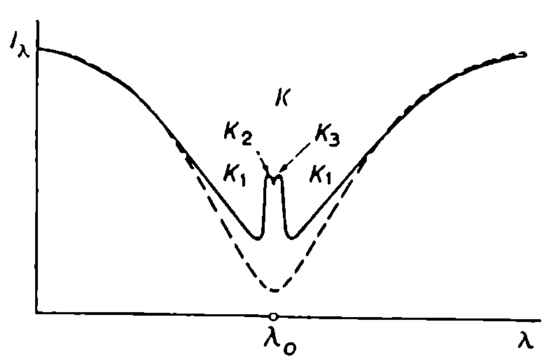

FIG. 57

The contour of the K line, in a spectrum taken at a point where there is a facula, has the form shown in Fig. 57. (The H line has the same kind of contour.) The bright core in the centre of the K line is called the K_2 line; the central depression in this core is called the K_3 line. Finally, the K_1 contour is the contour of the absorption K line itself, belonging to the facula proper. The lower, dashed, contour in Fig. 57 is the K line contour in the adjoining undisturbed parts of the photosphere.

It follows immediately from Figs. 51 and 57 that, in the central parts of the H, K and H_α lines, the faculae must in fact be brighter than the neighbouring regions of the undisturbed photosphere. Since this excess radiation (excess over the ordinary radiation from the undisturbed regions) is due to the chromospheric layers, and the regions where it is found coincide in general with the faculae, the corresponding bright regions seen on spectroheliograms are called **chromospheric faculae.** They are also often called **flocculi** or **facular plages.**

We may note the following important circumstance in connection with the interpretation of spectroheliograms in the H, K and H_α lines. Observations have shown that, the closer to the centre of the line the secondary slit of the spectroheliograph is adjusted, the higher are the layers of the chromosphere whose radiation is received. From the physical point of view, this is easily explained if we recall that the absorption coefficient in the line increases towards the centre. Hence the optical thickness of the chromospheric gases is greatest in the centre of the line, and no radiation reaches us directly from the deep layers of the chromosphere.

The main questions which arise in the interpretation of spectroheliograms in H, K and H_α are the following. What is the source of the excess energy radiated by the chromospheric gases in the central parts of these lines ? What are the causes of the broadening of the lines of this additional chromospheric radiation ? What is the origin of the central minimum K_3 in the bright core K_2 ? How are the flocculi related to the faculae lying beneath them ? Besides these questions, there are others, e. g. that of the origin of the hydrogen vortices round sunspots, and so on. However, we shall not concern ourselves with these latter questions here.

Let us begin by considering the source of the excess energy radiated by the flocculi. Here there are two possibilities: (1) the energy radiated in the flocculi arises as a result of fluorescence processes; (2) the chromospheric flocculi are regions which are hotter than the neighbouring parts of the undisturbed chromosphere.

The foundations of the theory of the former mechanism were developed by V. A. AMBARTSUMYAN. He first called attention to the fact that, when Q in formula (14.29) is fairly large, there is a range of values of η_ν close to the centre of the line such that, while the inequalities $\eta_\nu \gg 1$ and $\eta_\nu \varepsilon \ll 1$ are satisfied, the inequality $\eta_\nu \varepsilon Q \gg 1$ holds, so that r_ν takes the form

$$r_\nu \approx \varepsilon\, \eta_\nu\, Q \, \frac{\frac{4}{3}\sqrt{3}}{\sqrt{\eta_\nu} + \frac{2}{3}\sqrt{3}} \, \frac{B_\nu(T_0)}{B_\nu(T_E)} \,. \tag{21.26}$$

It follows from this expression that r_ν increases with η_ν, i. e. increases continually as we approach the centre of the line. In other words, we have an emission line in the centre of the absorption line.

A. Unsöld has indicated the second possibility, which presupposes a temperature inversion in the chromosphere. This hypothesis has recently been discussed by V. A. Krat [70] and others.

It is evident that the two mechanisms mentioned are not mutually exclusive. However, in the study of the physical properties of floccular regions, the primary problem is to decide which of these mechanisms is the principal one in each particular case. The following calculations are due to É. R. Mustel' [103].

In order to make the problem more definite, we propose the question: which is the main process governing the emission in the H and K lines and in the lines of the Balmer series in the spectra of flocculi: recombinations or exciting collisions?

Let us first compare the emission in the K and H_α lines, supposing that in each case the principal mechanism governing the radiation is recombination to the upper level. According to (5.48), the ratio of the numbers of captures to the corresponding levels is

$$A = \frac{[N_{f\to 2}]_{\mathrm{K}}}{[N_{f\to 3}]_{\mathrm{H}_\alpha}} \approx \frac{N(\mathrm{Ca\,III})}{N(\mathrm{H\,II})} \cdot \frac{[\phi(T_\varepsilon)]_{\mathrm{K}}}{[\phi(T_\varepsilon)]_{\mathrm{H}_\alpha}}, \tag{21.27}$$

where $N(\mathrm{Ca\,III})$ is the number of doubly ionised calcium atoms above 1 cm^2 of the base of the flocculus, and $N(\mathrm{H\,II})$ is the corresponding number of protons. The function $\phi(T_\varepsilon)$ is the probability of photo-recombination, which depends only on the kinetic temperature of the electron gas:

$$n_e\,\phi(T_\varepsilon) = \int_0^\infty \beta_\nu\, v\, \mathrm{d}n_e\,, \tag{21.28}$$

where β_ν is the effective recombination cross-section, determined by formula (5.55). Here we note that, in accordance with (5.52), the quantity n_e cancels on the right-hand side of (21.27).

The expected mean value of the electron temperature in flocculi lies between 5000° (the boundary temperature of the photosphere) and 10,000°. However, the function $\phi(T_\varepsilon)$ depends only very slightly on T_ε; hence we shall not commit any considerable error by taking $T_\varepsilon = 7500°$. We then find

$$[\phi(T_\varepsilon)]_{\mathrm{K}} \approx 1{\cdot}68 \times 10^{-13}\,, \quad [\phi(T_\varepsilon)]_{\mathrm{H}_\alpha} \approx 5{\cdot}0 \times 10^{-14}\,.$$

In other words, $A \approx 3{\cdot}4\,\{N(\mathrm{Ca\,III})/N(\mathrm{H\,II})\}$.

If the ionisation potentials of Ca II and H were the same, the ratio $N(\mathrm{Ca\,III})/N(\mathrm{H\,II})$ in the flocculus would be approximately equal to the relative content of Ca and H, i. e. of the order of 2×10^{-6} to 5×10^{-6}. (The number of neutral calcium atoms in the chromosphere, and therefore

in the flocculus, is negligible in comparison with the number of Ca II atoms.) In reality, this potential for Ca II is less by $\Delta\chi = 1{\cdot}7$ eV than for hydrogen, and consequently Ca II should be more strongly ionised than hydrogen. However, it is clear that, in the case considered, the degree of ionisation of Ca II cannot exceed that of hydrogen by more than one, or at the very most two, orders of magnitude, on account of the small difference $\Delta\chi = 1{\cdot}7$ eV. In this case the ratio A will be about 10^{-3}. Consequently, if the radiation in the K line were determined by recombination processes, the brightness of the H_α line would exceed that of the K line by three orders of magnitude; this is contrary to observation, since the brightnesses of the two lines are of the same order of magnitude. Thus *the emission in the* K_2 *line in flocculi must be determined by the collisions of* Ca II *atoms with electrons.* The fact that we are not here concerned with recombinations is confirmed by other factors also. For example, calculations show [106] that, if the radiation in the lines H_2 and K_2 in the spectra of flocculi were due to recombinations, the lines 3706 and 3737 of Ca II in these spectra would be only a little less bright than the H_2 and K_2 lines, which is completely contradictory to observation.

Let us make an approximate estimate of the temperature conditions which are necessary in order to produce the observed emission in the K_2 (or H_2) line. In a state of equilibrium, the number of exciting collisions $N_{e,1\to2}$ must be equal to the number of quanta $N_{2\to1}$ leaving the same volume. Let the numbers $N_{e,1\to2}$ and $N_{2\to1}$ be referred to a column above 1 cm² of the base of the flocculus. Then, by the general formula (8.21), we have

$$N_{e,1\to2} = N_1(\text{Ca II})\, n_e\, \bar{q}_{12} \sqrt{\left(\frac{8\,kT_e}{\pi\, m_e}\right)} \left(1 + \frac{\varepsilon_{12}}{kT_e}\right) e^{-\varepsilon_{12}/kT_e}\,, \quad (21.29)$$

where N_1(Ca II) is the number of Ca II atoms in the ground state, also above 1 cm² of the base of the flocculus, $\bar{q}_{12}$ is the mean effective cross-section for exciting collisions, which for the present we take as 10^{-16} cm², and finally ε_{12} is the excitation energy. We can determine the value of $N_{2\to1}$ directly from observations. These show that the intensity of emission in the K_2 line depends only slightly on the angle θ. In other words, this intensity is almost constant over the whole disc:

$$N_{2\to1} \approx 4\,\pi \int I_\lambda \frac{d\lambda}{h\nu}\,, \quad (21.30)$$

where the integration is extended over the whole of the K_2 (or H_2) line.

To determine N_1(Ca II) from the measured contour of the K_2 or H_2 line, we can use formula (20.3), which, however, requires some modification. Formula (20.3), if $P_\nu = j_\nu/s_\nu$ is constant with depth, takes the form

$$I_\nu = P_\nu(1 - e^{-\tau_\nu})\,, \quad (21.31)$$

and is applicable in the case where the observer receives radiation only from the emitting layer which he is examining. If, however, behind the emitting, and of course *absorbing*, layer considered, there lies another medium emitting with intensity I_ν^*, the total intensity recorded in the frequency ν will clearly be

$$I_\nu = I_\nu^* \, e^{-\tau_\nu} + \bar{P}_\nu (1 - e^{-\tau_\nu}) \,, \tag{21.32}$$

where the exponential in the first term on the right-hand side of (21.32) takes into account the attenuation of the beam of intensity I_ν^* by its passage through the emitting layer under consideration.

If the absorption coefficient s_ν is independent of depth, formulae (21.31) and (21.32) take the form

$$I_\nu = \bar{P}_\nu (1 - e^{-Ns_\nu}) \,, \tag{21.31 a}$$

$$I_\nu = I_\nu^* \, e^{-Ns_\nu} + \bar{P}_\nu (1 - e^{-Ns_\nu}) \,. \tag{21.32 a}$$

For the H_2 and K_2 lines in the spectra of flocculi, the background of intensity I_ν^* will evidently be produced by the central parts of the H and K absorption lines in the spectra of the faculae. The corresponding values of I_ν^* in the lines H_2 and K_2 can be obtained by extrapolating the contours of these absorption lines (H and K) to the centres of the lines.

The absorption coefficient $\bar{s}_\nu$ for the K_2 and H_2 lines is determined mainly by the Doppler effect. Here two limiting cases must be considered: (1) the emission in the K_2 and H_2 lines is produced predominantly in the lower layers of the chromosphere; (2) it is produced in the higher layers of the chromosphere. In the former case, turbulence may be neglected, and the width $\Delta \lambda_D$ is determined entirely by thermal motions. In the latter case we must also take account of turbulent motions, whose velocities, as we have said above, are of the order of 15 km/sec.

In the former case the fitting of the theoretical formula (21.32) to the observations gives a value $N_1(\text{Ca II}) \approx 10^{16}$ to 2×10^{16}. In the latter case we obtain $N_1(\text{Ca II}) \approx 2 \times 10^{13}$.

As a maximum value of n_e we can take 5×10^{11}. A value of $n_e \approx 10^{12}$ corresponds to faint chromospheric flares, and $n_e \approx 3 \times 10^{12}$ to strong ones.

Next, T_e in the flocculi cannot appreciably exceed 7500°. An increase in the temperature of the chromosphere (in its lower and middle layers) to 10,000° would lead (especially near the maximum of solar activity, when the area occupied by flocculi is sometimes only one order of magnitude less than the total area of the solar disc) to an extremely large increase in the ionisation of the ionosphere, a sharp rise in the radio emission, and so on. Hence we shall for the present take $T_e = 7500°$.

Using the values mentioned, we find that for $N_1(\text{Ca II}) = 2 \times 10^{16}$, $N_{c,\,1\to 2} = 2{\cdot}5 \times 10^{18}$, and for $N_1(\text{Ca II}) = 2 \times 10^{13}$ the value is $2{\cdot}5 \times 10^{15}$. On the other hand, for an average flocculus with a central intensity of the K_2 and H_2 lines of about 0·3, we find for $N_{2\to 1}$, using (21.30), the value 10^{18}. It follows from a comparison of these results that, firstly, the calcium flocculi lie in the lower layers of the chromosphere, where the turbulence is slight (it is easy to show that the same conclusion would hold even if we assumed an infinitely high temperature), and secondly, a temperature $T_\varepsilon = 7500°$ is wholly sufficient to produce the necessary emission in the K_2 and H_2 lines.

The first of these conclusions is in accordance with direct measurements, which give a height of 1300 km for the position of the K_2 flocculi. As regards the second conclusion, the existing data indicate that, for the alkaline earth elements, the group to which the Ca II atoms belong, the value of $\bar{q}$ for resonance lines is considerably greater than the ordinary cross-section given by the kinetic theory of gases, which is of the order of 10^{-16} cm^2; it may reach even 10^{-15} cm^2. Recent calculations by J. T. JEFFERIES [64] confirm this. They give, for energies of about 4 eV, a cross-section of $8{\cdot}7 \times 10^{-16}$ cm^2 for the K line and $5{\cdot}8 \times 10^{-16}$ cm^2 for the H line. In this case, the temperature T_ε which is necessary to produce the observed emission in the K_2 line is considerably decreased, to values of 5000 to 6000°. These temperatures are, according to a number of arguments, much closer to reality than one of 7500°.

It should be borne in mind that all the estimates just given, which relate to the mechanism by which calcium flocculi emit in the lines H_2 and K_2, refer to the *total* energy in these lines. It is clear that this must be the starting-point. merely because of the law of conservation of energy. However, if we construct the equation of transfer for frequencies within the H_2 or K_2 line and solve it, taking account only of collisions and using the above values for n_e and T_ε, the *intensity* I_ν so found is considerably less than the observed value. It can hardly be doubted that this is due to the neglect of non-coherent processes, which occur because of the thermal motions of the atoms in the chromosphere (see Chapter 14). S. MIYAMOTO [91] has shown that such processes should be very important in this problem.

Let us now consider the question of the origin of the central lines K_3 and H_3. Their presence may be due to a rapid decrease in the value of P_ν outwards in the flocculus. It is easy to see that this will in fact lead to the appearance of a dark core, by direct calculations based on formula (20.3). This conclusion is in agreement with the results of M. WALDMEIER already given (see Chapter 19), who found that in the faculae, i. e. immediately below the flocculi, the value of P_ν also decreases outwards, beginning from a mean optical depth $\tau \approx 0{\cdot}6$. Thus the flocculi and the

faculae are related not only by being situated close to each other, but also by the general distribution of temperature in them.

It should be noticed that this decrease in P_ν, and therefore in T_ε, as we move outwards cannot continue to an indefinite height. Even in the middle chromosphere, it must give way to an increase in T_ε as we approach the corona, as in other regions of the Sun.

MIYAMOTO [91] gives another possible explanation of the origin of the H_3 and K_3 lines, which involves an allowance for non-coherent processes (see above). The physical nature of the process is as follows. Since the chromosphere is very opaque in the cores of the H and K lines, the quanta produced by exciting electron collisions within the Doppler core will emerge from the medium (as a result of non-coherent redistribution of the radiation in frequency) in the more transparent wings of the line concerned. The central "opaque" part of the line has, according to MIYAMOTO's calculations, a fairly low intensity, and he identifies this with the H_3 and K_3 lines. The same calculations account for the great width of the H_2 and K_2 lines.

The following remark should be made in connection with this last result. On taking upper limiting values which seem reasonable for n_e and T_ε, and using formula (21.29), and comparing $N_{e,1\to2}$ with the value of $N_{2\to1}$ found from observation, we again get a large value of N_1(Ca II), of the order of 10^{15} to 10^{16} cm^{-2}, and therefore a great width for the H_2 and K_2 lines. Hence the conclusion that the H_2 and K_2 lines are very wide can be drawn whether or not the emission process is coherent.

The arguments of MIYAMOTO given above are very convincing. However, they meet with difficulties when we try to explain the contours of the bright H and K lines in the spectra of spots. It is well known that above sunspots the bright central reversals H_2 and K_2 are much narrower and do not show the dark H_3 and K_3.

The explanation of these two facts on MIYAMOTO's theory does not seem possible. On the other hand, they *can* be explained as follows [107]. Firstly, the narrowing of the bright central reversals (as we go from the adjoining flocculus to the spot itself) cannot be the result of a decrease in the number of *emitting* atoms above the spot, since observation shows that the width of the bright H_2 and K_2 lines in flocculi and of the very faint H_2 and K_2 lines in undisturbed parts of the Sun's disc is about the same. Hence this width is determined, not by the number of emitting atoms, but by the number N_1(Ca II) of *absorbing* Ca II atoms in the chromosphere (the effect of self-absorption; see formula (21.31a)). We can therefore conclude that there are fewer atoms of Ca II (and of other elements also) in the chromosphere above spots than there are above neighbouring regions of the Sun's surface. This is in agreement with

many direct observations of the chromosphere both outside and during eclipse, which indicate a considerable decrease in the height of the chromosphere above spots.

This decrease might be explained by the outflow of matter from a spot in its lower chromospheric layers (the Evershed effect for faint lines). Moreover, the lowering of the chromosphere over a spot could also explain the inflow of matter into a spot in the higher layers of the chromosphere (the Evershed effect for strong lines), since in this case the density of the chromosphere above the spot would be less than in the adjoining layers of the chromosphere.

The absence of the H_3 and K_3 lines in the bright central reversals above spots could be explained by the fact that the temperature of the lower layers of the chromosphere, which is of the order of 5000°, is close to the temperature of the spot (4500°). This means that, in the chromospheric layers which form the bright central reversals in the H and K lines, the kinetic temperature does not vary greatly with height, and may even increase upwards. This has the consequence that the value of P_ν in these layers either is constant or increases outwards, and so the H_3 and K_3 lines do not appear.

Let us now consider what determines the emission in the H_α line in the spectra of flocculi (É. R. Mustel' [105]). This question is much more complicated than the corresponding one for the H_2 and K_2 lines of Ca II. In fact, to compare the parts played by recombinations and by exciting collisions, we need to know N_1 (H), the number of neutral hydrogen atoms in the ground state above 1 cm² of the base of the flocculus. (The number of collisions which excite from the second level is very small.) The determination of this quantity, however, is soon seen to involve great difficulties. If the lower layers of the flocculus coincide approximately with those of the undisturbed chromosphere, calculations show that recombinations and exciting collisions at $T_\varepsilon \approx 7500°$ are of approximately equal importance in producing the emission in H_α. However, we have seen that T_ε in the flocculi is apparently less than 7500°. In this case recombinations should play a more important part than exciting collisions.

These considerations seem to account for the very different appearance of spectroheliograms taken in the lines H and K of Ca II, as compared with those taken in the lines of the Balmer series (particularly H_α). In view of the importance of this question, we shall discuss it in more detail.

Let us first consider those parts of the Sun's surface in which there is no bright floccular radiation. For these regions, the kinetic temperature of the lower layers of the chromosphere (where most of the radiation in the central parts of the Fraunhofer lines H and K of Ca II and H_α originates) is of the order of 5000° to 6000°; and the recombination

processes in H_α probably predominate over the processes of excitation of hydrogen atoms by electrons. On the other hand, we have seen that even a low electron temperature of the order of 5000° to 6000° is sufficient to excite the atoms of Ca II. Hence the appearance of the spectroheliograms ought to be entirely different for calcium and for hydrogen lines, and this is confirmed by observation. It is well known that one of the chief properties of calcium spectroheliograms is their "spotty" appearance (even for the undisturbed regions of the Sun), with no sign of vortical or filamentary features. The individual bright spots and lines are the most usual characteristic of calcium spectroheliograms. On the other hand, the distinguishing feature of hydrogen spectroheliograms (particularly those taken in H_α) is that they show a turbulent and somewhat filamentary appearance. These features are most distinctly seen in the neighbourhood of sunspots, where, as is well known, spiral vortices are observed.

On the basis of the ideas discussed previously, the facts just enumerated can be explained as follows. Hydrogen is the predominating element in the solar envelopes. This enables us to suppose that the turbulent (and filamentary) structure of the spectroheliograms taken in the H_α line is caused by inhomogeneities in the mean density distribution of chromospheric matter over the disc. Since, for hydrogen, the intensity of recombination radiation is proportional to n_e^2, even small oscillations in this mean density will have a considerable effect on the distribution of the intensity of radiation*. Hence, if the distribution of matter near the spot is vortical, the radiation field will be so too. On the other hand, the variations in the intensity in the K_2 and H_2 lines over the disc depend much less on variations in the density of matter; they are chiefly affected by changes in the electron temperature T_ε (see formula (21.29)). Hence it will not be easy here to detect the vortical structure.

The fact that in the undisturbed regions of the Sun the recombination process (whatever the value of T_ε) is responsible *quantitatively* for the observed radiation can be proved by direct calculations [**105**]. Furthermore, it can be shown that the observed radiation in the central part of the Fraunhofer line H_α originates in the lower layers of the chromosphere.

Let us now turn to those regions of the Sun's surface where bright floccular radiation is observed. The most important difference here between calcium and hydrogen spectroheliograms is that the area occupied by calcium flocculi is much greater than that occupied by hydrogen flocculi. This is undoubtedly due to the "easier" excitation of chromospheric calcium. We have seen that a quite small increase of the temperature T_ε in the chromosphere, to 6000°, is enough to produce a con-

* Of course, the variations in the degree of ionisation of hydrogen must be of some importance also.

siderable amount of radiation in the H and K lines. On the other hand, such a temperature would be entirely inadequate to excite hydrogen or to ionise it to any noticeable extent.

The problem of the mechanism which produces the radiation of solar regions where bright floccular radiation is observed in H_α (and, of course, in H and K of Ca II) is not yet resolved. It seems quite evident, however, that, in the regions of the Sun's disc where bright hydrogen flocculi are observed, the kinetic temperature of the chromosphere must be fairly high, or else some very efficient ionisation mechanism must be in operation. It must be remembered that both the ionisation potential and the excitation potential from the ground level are quite high for hydrogen atoms. For this reason, we should expect that helium flocculi would resemble hydrogen ones in shape and size, and not calcium ones; as we have said, a relatively slight increase in T_ε is sufficient to form calcium flocculi.

These latter considerations [104] regarding helium flocculi are fully confirmed by observation. According to the results of M. and L. D'AZAMBUJA, there is a very close similarity in the appearance of spectroheliograms taken in the line λ 10,830 Å of He I and in the H_α line, the only difference being that the flocculi are bright in H_α and dark in λ 10,830 Å.

However, A. A. NIKITIN [115] has shown that the appearance of this helium absorption line λ 10,830 Å in the spectrum of disturbed regions of the Sun's disc can be explained by supposing that it is formed on account of an accumulation of helium atoms in the 2^3S state, as a result of recombinations to all the triplet levels.

The exact correspondence between hydrogen and helium flocculi shows, *inter alia*, that the state of excitation and ionisation in regions where there is bright floccular radiation must be fairly high — higher than would correspond to a temperature of $T_\varepsilon = 7500°$. The latter temperature would be quite inadequate to bring about any noticeable ionisation of helium.

On the other hand, the excitation state cannot be very high, since we should then notice the bright line λ 5876 of He I.

Let us now turn to the last question, that of the origin of the increased kinetic temperature in flocculi and of the anomalous ionisation of hydrogen and helium. It might be supposed that these phenomena are related to the anomalously strong coronal radiation above the faculae (and flocculi). The anomalously strong ultra-violet radiation of the corona ionises the chromospheric hydrogen, and this in turn leads to a raising of the kinetic temperature of the electrons. However, we must bear in mind that the faculae lying beneath the flocculi are physically connected with them, and the above "ultra-violet" theory cannot be regarded as corresponding to reality in the case of the faculae, since

the optical thickness of the latter is very large in the ultra-violet region of the spectrum. Furthermore, this mechanism cannot easily explain the fact that the regions of the corona where the intensity of monochromatic coronal radiation is enhanced usually occupy positions on the Sun of quite large vertical and lateral extent. Hence we should expect that the area occupied by (e. g.) helium and hydrogen flocculi would be even greater than that occupied by calcium flocculi. This is completely contradictory to observation. The regions which are bright in H_α and correspondingly dark in λ 10,830 Å of He I often take the form of narrow curved strips or separate points; in some cases, the bright radiation in H_α follows the arms of vortices, and so on.

Contrary to the above hypothesis, it may be assumed that the anomalous conditions in flocculi, the anomalous temperature gradient in faculae (where there is neither radiative nor convective equilibrium), and the anomalous ionisation of Ca II and of other elements in faculae, are all the result of the existence of very hot coronal matter in these regions, which is found alongside the low-temperature material of the chromosphere. The effect of the coronal matter on this latter material should appear most strongly for high excitation and ionisation energies, and this is in fact the case. Such a coexistence of two different "phases" of matter is observed, for instance, in prominences (see Chapter 20). This hypothesis is in accordance with the fact that the monochromatic radiation of the corona is especially strong directly above facular regions.

8. **Chromospheric flares.** To conclude the present chapter, we shall briefly discuss the subject of chromospheric flares. A **chromospheric flare** is a sudden increase in the brightness of an individual region of the solar surface, which is most frequently observed in the H_α line, but appears also in some other lines of the solar spectrum, for example, in the lines of He I, Ca II, Fe I, Fe II, Si II, Ti II, Sr II, Sc II, the D_1 and D_2 lines of Na I, etc. The D_3 line of He I is sometimes observed in absorption also. Very rarely, an intensification of the continuous spectrum also is observed in the region of a flare.

Flares most frequently appear in the region occupied by bright flocculi, and, as a rule, near sunspots. Sometimes, though very rarely, they appear in a region where no other feature was previously seen. The area occupied by a flare varies in different cases, sometimes reaching 0·005 of the surface of the Sun's disc.

Flares, if observed at the limb of the disc, are similar in appearance to low but very bright prominences. The matter in these "prominences" may lie either inside or outside the chromosphere. In some cases the flares are situated even deeper, in the photosphere (F. Ellerman). These flares appear in the form of emission *in the wings only* of the H_α line. The

bright emission in the central parts of the line, where the absorption coefficient is large, is in this case stopped by the higher-lying atoms of the "reversing layer".

The observed area of a chromospheric flare lying at some distance from the centre of the disc is composed of two parts. The first part is the projection of the "lateral surface" of the flare, while the other is the projection of the "top" of the flare. An analysis of measured areas of flares observed at various distances from the centre of the disc [180] gives a law of the form $e^{-\alpha h}$, with $\alpha^{-1} = 20{,}300$ km, for the distribution of the altitudes of flares.

The cinematography of flares at the Sun's limb shows that they have a very great variety of forms. Some characteristic instances have been described by H. W. DODSON and R. R. MCMATH [37], and also by A. B. SEVERNYĬ [131].

The brightness of a flare in H_α usually increases rapidly and then decreases more slowly. An investigation of the development of chromospheric flares by A. B. SEVERNYĬ and E. F. SHAPOSHNIKOVA [135], using cinematography in H_α, shows that flares, including small ones, appear on the average at the rate of one every 7 hours during the lifetime of a spot group, even during years of medium activity (1951 to 1953). More than 30 % of all flares are seen to be in motion (in a "sideways" or horizontal direction), in the same manner as sunspot prominences (with velocities of up to 300 km/sec); the remainder show a more or less simultaneous expansion of their area in all directions. As a rule, the brightness of a flare increases together with its area, and the more rapid the expansion, the greater the maximum brightness reached. The "lifetime" of a flare (the half-width of the brightness curve) is the greater, the greater the area of the flare.

A number of investigators have also established that the brightness of a flare and the width of the H_α line vary in a correlated manner.

The contour of the H_α line in the spectrum of a flare is shown in Fig. 51. This figure represents a flare of moderate brightness. In the case of intense flares, the brightness of the central peak may exceed that of the adjoining continuous spectrum by a factor of 1·5, or sometimes 2.

The study of the spectra of flares has shown that the bright centre in the H_α and other lines is not, as a rule, displaced relative to the normal position of the lines concerned. In other words, the luminous matter in a flare has no appreciable velocity in the line of sight. It is possible that in some cases a motion of the matter above the photosphere takes place in the *earliest* period of the existence of the flare, but the brevity of this period causes great observational difficulties*.

* One such case is described by DODSON and MCMATH [37].

A prominence (or several prominences) is usually ejected from the region occupied by a fairly bright flare, and this prominence belongs to the surge class. The matter, having been ejected, moves upwards with high velocity to a certain height, and then returns downwards along the same trajectory. In projection on the disc, this matter is visible in H_α light as a dark spot on the brighter background of the disc. The ejection of matter is, however, sometimes not followed by its return.

The appearance of a chromospheric flare on the Sun is accompanied by bursts of radio emission from the region of the flare. A study of this radio emission at decimetre and metre wavelengths has revealed a number of interesting features [36]. Moreover, motions of the radio source from the flare out into the corona have been discovered. It is not yet certain whether this motion is related to the surge prominence.

Bright chromospheric flares are usually accompanied by a sharp deterioration in the reception of short radio waves. Moreover, if the flare is in the central part of the Sun's disc (not more than 45° from the centre of the disc), and if it is sufficiently bright, then a perturbation of the Earth's magnetic field and interference with radio communication are observed about 24 hours after the appearance of the flare.

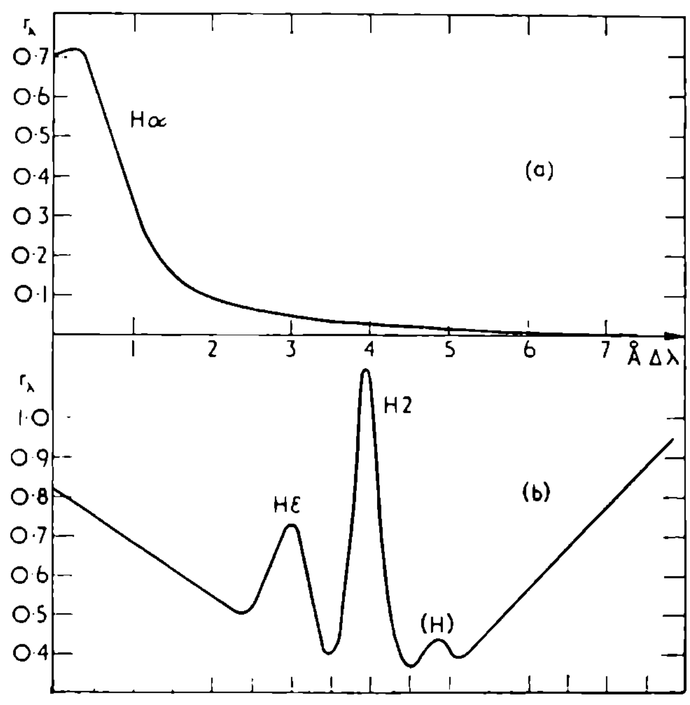

Fig. 58

Finally, it should be mentioned that chromospheric flares are sometimes accompanied by an increase in the intensity of cosmic rays.

After these short introductory remarks, let us consider some problems of the physics of chromospheric flares. We first discuss the lines of the Balmer series in the spectra of flares. Observations show that, in the spectra of fairly intense flares, practically all the lines of the Balmer series exhibit bright centres. This is most markedly apparent in the H_α line. For very intense flares, this reversal in H_α extends over a large wavelength range, so that the total width of the "wings" of the reversal may reach 16 Å. Fig. 58(a) shows the emission contour of the H_α line for the large chromospheric flare of 5 August 1949, obtained

by A. B. SEVERNYĬ and É. R. MUSTEL' from spectra taken by them at the Crimean Astrophysical Observatory. This contour (the red wing) was obtained by subtracting the intensity of radiation in the H_α line in the spectrum of the undisturbed parts of the solar surface from the intensity observed in the flare. (This contour gives only the "extra" radiation due to the flare. In a theoretical analysis, we have to start from the contour actually observed and apply formula (21.32).) We see that the wings of the H_α line in emission extend to 6 or 7 Å from the centre of the line. The H_β, H_γ, ... lines in the spectra of flares are considerably narrower. Fig. 58(b) shows the contours of the H_ε and H_2 emission lines in the spectrum of the same flare. The half-width of H_ε is here only about 2 or 3 Å. The letter (H) marks the radiation in the H line due to the surge prominence ejected from the flare.

Two main questions arise regarding these results: firstly, what brings about so great a broadening of the H_α emission line in flares, and secondly, why the H_β, H_γ, ... emission lines are narrower than the H_α line. É. R. MUSTEL' and A. B. SEVERNYĬ [111] have shown that the wings of the H_α emission line are due to ordinary processes of radiation damping, whereas the broadening of the H_γ, H_δ, H_ε, ... lines is caused by the Doppler and Stark effects. This hypothesis explains the contours of all these lines in the spectrum of the flare (apart from the centres of the H_γ, H_δ and H_ε lines, where the hypothesis that P_ν is constant with depth involves considerable errors). It was necessary here to take account of self-absorption in the central parts of the lines; it is particularly strong for the H_α line.

The theoretical contours were calculated from formula (21.31a).* In the case of damping and the Doppler effect (for H_α), the coefficient s_ν is determined by formula (11.39), in which the quantity $\Delta\nu_D$ enters as a parameter.

On the other hand, for the H_γ, H_δ and H_ε lines the value of s_ν is determined by the simultaneous action of the Doppler and Stark effects.† Thus, we have to account for all the emission contours of H_α, H_β, H_γ, H_δ, H_ε, ... by choosing three parameters: N_2, n_e and $\Delta\nu_D$, where N_2 is the number of hydrogen atoms in the second excitation state above 1 cm² of the base of the flare. For the flare of 5 August 1949 it was found that for the brightest parts $N_2 = 1{\cdot}3 \times 10^{16}$, $n_e \approx 3 \times 10^{12}$, and $\Delta\nu_D$ corresponded to a temperature $T \approx 11{,}000°$.

The above-mentioned conclusion that the broad wings of the H_α line in the spectra of chromospheric flares are caused by radiation damping, and that the optical thickness of flares in the central parts of this line

* It would be more correct to start from formula (21.32a), but no great error arises in the numerical results through not doing so.

† For a table of values of s_ν for the lines H_α to H_δ, see S. VERWEY [171].

must be very large, has been confirmed by recent work by L. GOLDBERG, H. W. DODSON and E. A. MÜLLER [44]. Furthermore, these authors find that the increase, derived from observation, in the width of the emission line H_α with distance from the centre of the Sun's disc indicates that, when the flare is at the limb, the optical path in it is large (i. e. the number of atoms in the line of sight is large). This means that the height of a flare is usually less than its linear dimensions in the "horizontal" directions.

Knowing the absolute intensity of radiation in the H_α, H_β, ... lines, we can also determine N_3, N_4, ... i. e. the number of emitting atoms in the respective quantum states. Let us consider, as an example, the H_α line.

According to (20.1), the integrated intensity in the H_α line is determined by

$$I_{23} = N_3 A_{32} h\nu_{23}/4\pi , \tag{21.33}$$

where self-absorption is supposed absent. If the emission (and absorption) processes are determined by radiation damping and the Doppler effect, the law which gives the frequency distribution of the emitted energy I_{23} is given by formula (11.39). According to what we have said concerning this formula, the required law for the case $a \ll 1$ has the form

$$\phi(\nu) = \frac{1}{\sqrt{\pi}\,\Delta\nu_D}\left[e^{-(\Delta\nu/\Delta\nu_D)^2} + \frac{\delta_{ik}}{\sqrt{\pi}}\,\frac{\Delta\nu_D}{(\nu-\nu_0)^2}\right], \tag{21.34}$$

where

$$\int_0^\infty \phi(\nu)\,\mathrm{d}\nu = 1 . \tag{21.35}$$

Thus the value of I_ν for any frequency ν is, for transitions $3 \to 2$,

$$I_\nu = \frac{N_3 A_{32} h\nu_{23}}{4\pi^{3/2}\Delta\nu_D}\left[e^{-(\Delta\nu/\Delta\nu_D)^2} + \frac{\delta_{23}}{\sqrt{\pi}}\,\frac{\Delta\nu_D}{(\nu-\nu_0)^2}\right]. \tag{21.36}$$

For the extreme wings, the first term is negligible. In this case, I_ν has the form

$$I_\nu = \frac{N_3 A_{32} h\nu_{23}}{4\pi^2}\,\frac{\delta_{23}}{(\nu-\nu_0)^2} . \tag{21.37}$$

Consequently, having found I_ν in the wings from observation, we can determine N_3 also. For the flare of 5 August 1949, the value so found was $N_3 \approx 3 \times 10^{14}$, which is two or three orders of magnitude greater than for the ordinary chromosphere.

A comparison of the energy radiated in the H_α line in the flare and in the chromosphere shows that the increase in this energy in the flare is a direct consequence of the increase in n_e, which by (21.21) brings about an increase in the number of recombination processes. Thus, in a flare we are concerned simply with an increased ionisation of hydrogen.

Since the value of N_2 is very large in bright flares, these radiate a very great amount of energy in the Lyman α line. It is possible that this energy is sufficient to bring about the observed changes in the state of the Earth's ionosphere. However, there are difficulties to be met here, since, according to the most recent results, the solar radiation in the Lyman α line, in its path to the D layer, is strongly absorbed by the higher layers of the Earth's atmosphere.

The great width and intensity of the Lyman α line might explain the ejection of hydrogen atoms from flares (with velocities of up to 1000 or 3000 km/sec) under the action of radiation pressure. The existence of such currents of hydrogen atoms above flares follows, for instance, from the observed asymmetry in the emission contour of the H_α line in flares (M. A. Ellison [40]). Moreover, the existence of these currents should be expected on account of the occurrence of geomagnetic disturbances of a corpuscular nature, observed approximately 24 hours after the appearance of a bright chromospheric flare on the Sun. However, the idea of radiation pressure meets with serious difficulties, since the hydrogen atoms, in their motion away from the Sun, should be ionised by the Sun's coronal radiation, which would considerably reduce the radiation pressure.

We have here discussed the emission line spectrum of hydrogen in flares. It is found that the contours of the H and K emission lines of ionised calcium in flare spectra can also be interpreted as the result of Doppler broadening and radiation damping [134]. For the flare of 13 June 1950, the following parameters were obtained from a comparison of theory and observation: $N_1(\text{Ca II}) = 7{\cdot}2 \times 10^{16}\ \text{cm}^{-2}$; for the H line, $N_2 = 0{\cdot}64 \times 10^{14}\ \text{cm}^{-2}$; for the K line, $N_2 = 1{\cdot}25 \times 10^{14}\ \text{cm}^{-2}$, where N_2 gives the number of atoms in the second excited state of Ca II. For the same flare it was found that the turbulent velocity v_t apparently could not exceed 7 to 10 km/sec.

The energy radiated by a flare in the ultra-violet region of the spectrum may ionise atoms in the reversing layer. This has been directly confirmed by the results of É. R. Mustel' and A. B. Severnyĭ [110].

We may make some remarks concerning the origin of the continuous spectrum observed in some of the brightest chromospheric flares. I. M. Gordon [49] thinks that here we are concerned with the radiation of relativistic electrons in magnetic fields, the presence of which near chromospheric flares he postulates. However, there are more natural

ways in which this phenomenon could be explained [108]. Firstly, this continuous emission may be due to *recombinations*. Calculation shows that, for very bright chromospheric flares, this mechanism may be sufficiently effective. Secondly, and more probably, the continuous radiation might be formed when the lower parts of the flare are in the photosphere. We have seen that chromospheric flares are situated at very different levels. It is therefore quite natural to suppose that some very bright flares occupy a fairly wide range of altitudes, including the photospheric layers. In this case an increase in the temperature of the photosphere by 150° or 200° would be sufficient to give the observed intensity of continuous radiation. This increase may be caused by the transfer of thermal energy from the flare to the neighbouring parts of the photosphere.

The question of the origin of flares is still unresolved. However, two possibilities may be mentioned: (1) a flare is a kind of electric discharge in the chromosphere, which results in an increase both in the electron temperature and in the ionisation of hydrogen; (2) a flare is a sudden transfer of coronal matter from the lower layers into the chromosphere and an increased ionisation of the chromosphere in the neighbourhood of the flare (the same is true of flocculi). It is as yet difficult to say which of these two hypotheses corresponds more closely to reality. Further investigations are necessary before we can do so.

Chapter 22. The corona and the radio emission of the Sun

1. **General description and spectra.** The outermost and most extensive part of the Sun's atmosphere, the **corona**, could for a long time be observed only during a total eclipse. This was due to the fact that the surface brightness of the corona is about a million times less than that of the photosphere. In photographs, the corona has a complex structure. It was shown by A. P. Ganskiĭ that the shape of the corona depends on the phase in the cycle of solar activity: at maximum, the corona surrounds the Sun's disc nearly uniformly, while at minimum it is markedly flattened at the poles. The total brightness of the corona, according to V. B. Nikonov and E. K. Nikonova [116], also varies with the phase. The corona is characterised by a ray structure: there are **polar rays**, which are similar in form to the lines of force near a magnetised sphere, intense **straight rays** above faculae, based on chromospheric spicules, and **rays above spots**, which diverge at their base towards the sides of the spot and approach again at greater heights. Prominences are usually surrounded by a system of cupola-like envelopes,

above which rise **helmet-shaped rays.** These in turn divide into **jets,** between which there is more rarefied matter. The structure of the corona has been investigated in detail by Soviet scientists [25, 174].

The colour of the corona is in general close to that of the Sun, although G. A. TIKHOV has discovered that the outer parts of the corona (other than the rays) are slightly reddened in comparison with the photosphere.

The spectrum of the inner parts of the corona is a continuous one, on which bright lines are superposed; 24 coronal lines are at present known. The distribution of energy in the continuous spectrum of the corona is very similar to that in the continuous spectrum of the Sun. Among the bright lines, the green line $\lambda = 5303$ Å, and in some regions the red line $\lambda = 6374$ Å, are particularly noticeable. These lines are so bright that B. LYOT in 1930 succeeded, by selecting them, in observing the inner parts of the corona outside eclipse. During the observations, careful steps were taken to diminish the scattered light of the Sun in the instrument, the coronagraph itself being located on a mountain top to minimise the trouble due to light scattered in the atmosphere.

Observations outside eclipse have shown that the green and red lines are strong in individual regions of the corona, which are called **green regions** and **red regions,** and are concentrated round the active regions of the Sun, their connection with spots and other objects on the surface of the Sun being well defined. At fairly high latitudes ($\phi > 60°$) the green line is seldom observed, and at the poles the red line vanishes also.

In the outer parts of the corona (at distances of 0·2 to 0·3 $R_\odot$ from the Sun's surface) the emission lines become faint in comparison with the continuous spectrum, and absorption lines appear, which have approximately the same width as in the Sun's spectrum, but are much less deep. It has been found empirically that the spectrum of the corona can be regarded as a superposition of two spectra: a Fraunhofer component, in which the absorption lines have the same depth as in the Sun's spectrum, and a non-Fraunhofer component, in which there are no absorption lines at all.

The energy distribution in the non-Fraunhofer component of the spectrum is close to that in the continuous spectrum of the Sun, but exhibits some diminution in intensity in the violet region of the spectrum. The radiation of the corona is partly polarised. At the present time it is regarded as established with certainty that the non-Fraunhofer radiation of the corona is due to the scattering of the light of the photosphere by free electrons. It is known that a free electron scatters in a given direction a fraction $\mathrm{d}s_e$ of the radiation incident on it:

$$\mathrm{d}s_e = \frac{e^4}{2\,m^2 c^4}\,(1 + \cos^2\theta)\,\mathrm{d}\omega\,. \tag{22.1}$$

The total effective cross-section, or scattering coefficient, is

$$s_e = \int \mathrm{d}s_e = \frac{8\pi}{3}\left(\frac{e^2}{m\,c^2}\right)^2 , \tag{22.2}$$

so that s_e is independent of the wavelength. The scattered light is completely polarised if the angle by which the quantum deviates from its original direction (the scattering angle) $\theta = 90°$, and is unpolarised if $\theta = 0°$ or $180°$. For other values of θ, the scattered light is partly polarised. As we move away from the surface of the Sun, θ approaches $90°$, and the polarisation of the non-Fraunhofer component increases.

The corona is composed of ionised gases, which are as a whole neutral, since no considerable volume charge can exist, on account of the high conductivity of the Sun's atmosphere. The negative charges of the electrons exactly balance the positive charges of the ions, which are chiefly protons, since the atmosphere of the Sun consists mainly of hydrogen. We shall discuss below the chemical composition of the corona.

The absence of absorption lines in the spectrum of the inner corona is explained by the obliteration due to the large Doppler broadening when light is scattered by fast-moving electrons.

Each electron scatters radiation non-selectively if it is stationary relative to the source (the Sun) and to the observer. If the electron is in motion, the frequency of the scattered light will differ from that of the incident radiation by a quantity determined by the velocity of the electron relative to the Sun and to the observer. The velocity corresponding to the boundary temperature of the Sun, i. e. 5000°, is sufficient to obliterate faint lines, but G. A. SHAÏN [142] has shown that in the inner corona no traces are observed even of the very strong H and K lines of Ca II. It is easily calculated that an electron velocity dispersion corresponding to a kinetic temperature of over 500,000° is necessary to produce this amount of obliteration.

The Fraunhofer component of the radiation of the corona is due to the scattering of the Sun's radiation by small solid dust particles, which move in interplanetary space (H. C. VAN DE HULST, C. W. ALLEN). The scattering of the Sun's radiation by these particles causes the zodiacal light. In the immediate neighbourhood of the Sun, these particles would vaporise and they cannot exist there. Nevertheless, an increase in the brightness of the Fraunhofer component towards the Sun is observed (in the inner corona, the Fraunhofer component is not observed, despite its greater brightness, since the brightness of the electron component increases there to a still greater extent), and this is explained by the particular nature of the scattering. The indicatrix of scattering has a sharp maximum in the direction of the light incident on the dust particle, and hence we see light scattered by particles lying between the Sun

and the Earth. In order to "see" particles lying at a large angular distance from the Sun, we should have to detect light scattered through a large angle, and the effective cross-section for this kind of scattering is very small.

Such a theory gives a good explanation of the observed properties of the Fraunhofer component of the radiation of the corona, namely, the brightness' being independent of the position angle, the slow decrease in brightness with distance, the slight reddening compared with the colour of the Sun, and the absence of polarisation, and it is generally accepted at the present time. Thus only the non-Fraunhofer component is emitted by the corona itself, and in what follows we shall discuss only this component.

2. **The electron concentration.** It is natural to ask how the concentration of the electrons which give the observed brightness of the continuous spectrum of the corona can be determined. Unfortunately, there are few observations of the corona which are suitable for photometric reduction. S. Baumbach has studied photometric data from ten eclipses, and has obtained the brightness distribution in the corona (including the Fraunhofer component), averaged over all the observations and over position angles, as a function of ϱ, the distance from the centre of the disc in terms of the Sun's radius. This distribution is represented by the empirical formula

$$I(\varrho) = \frac{0\cdot0532}{\varrho^{5/2}} + \frac{1\cdot425}{\varrho^{7}} + \frac{2\cdot565}{\varrho^{17}}\,. \tag{22.3}$$

The unit of $I(\varrho)$ is a millionth part of the brightness of the centre of the Sun's disc. Since the energy distribution in the continuous spectra of the Sun and of the corona is very similar, the relation (22.3) holds for all frequencies in the visible region of the spectrum. The brightness of the corona depends very noticeably on the phase in the solar activity cycle and on the position angle, and hence (22.3) gives only an approximate mean value of $I(\varrho)$.

The surface brightness (intensity) $I(\varrho)$ and the emission coefficient* $j(r)$ are related by the integral equation

$$I(\varrho) = \int_{-\infty}^{\infty} j(r)\,\mathrm{d}y\,, \tag{22.4}$$

if the corona is assumed to be spherically symmetrical and scattering is isotropic (i. e. we neglect the dependence of $\mathrm{d}s_e$ on θ). Solving the

* $j(r)$ is referred, not to 1 cm³ as usual, but to a column oriented along the y axis, of length $R_\odot$ and cross-section 1 cm².

integral equation (22.4) in the same manner as (21.6), but taking into account the greater extent of the corona, we find for $j(r)$ the expression

$$j(r) = \frac{0{\cdot}0304}{r^{7/2}} + \frac{1{\cdot}452}{r^{8}} + \frac{4{\cdot}157}{r^{18}} \,. \tag{22.5}$$

On the other hand, $j(r)$ can be written in the form

$$j(r) = s_e\, R_\odot\, n_e(r) \int \frac{I_\odot\, \mathrm{d}\omega}{4\pi} \,. \tag{22.6}$$

The factor $R_\odot$ (the radius of the Sun in centimetres) appears on the right-hand side because we have used the Sun's radius as the unit of length of $\mathrm{d}y$ in (22.4). The integration must be over ϕ from 0 to 2π, and over the second spherical co-ordinate ψ from 0 to ψ_0, where ψ_0 is half the angle subtended by the Sun at a point P distant r from its centre. The expression (4.34), which takes account of the darkening to the limb of the Sun's disc:

$$I_\odot = I_\odot{}^0 (1 - u + u \cos\theta) \,, \tag{22.7}$$

is taken for the intensity of the incident radiation $I_\odot$. After a series of calculations, Baumbach found the following formula for $n_e(r)$:

$$n_e(r) = \frac{\dfrac{0{\cdot}0304}{r^{7/2}} + \dfrac{1{\cdot}45}{r^{8}} + \dfrac{4{\cdot}16}{r^{18}}}{R_\odot \dfrac{8\pi e^4}{3m^2c^4}\, \frac{1}{2}\, I_\odot{}^0 \left\{ (1-u)\left(1-\sqrt{\left[1-\frac{1}{r^2}\right]}\right) + \frac{1}{2}u\left[1-r\left(1-\frac{1}{r^2}\right)\log_e \sqrt{\left(\frac{r+1}{r-1}\right)}\right]\right\}} \,, \tag{22.8}$$

which, putting $u = 0{\cdot}8$ (for the photographic region of the spectrum), he replaced by the simpler interpolation formula

$$n_e(r) = 10^8 \left(\frac{0{\cdot}036}{r^{3/2}} + \frac{1{\cdot}55}{r^{6}} + \frac{2{\cdot}99}{r^{16}} \right) . \tag{22.9}$$

We see that the density in the corona decreases very slowly with distance.

Baumbach's data have been subjected to a series of corrections. H. C. van de Hulst has excluded the Fraunhofer component, which does not belong to the corona, from the total brightness of the corona, and has calculated $n_e(r)$ with the new value of $I(\varrho)$. In the outer corona, the difference between his results and the previous ones is quite considerable.

A. F. Bogorodskiĭ and N. A. Khinkulova [23] have carried out a calculation of $n_e(r)$, taking into account the fact that $\mathrm{d}s_e$ depends on θ. The calculation presents great mathematical difficulties, and we shall not explain it in detail.

An element of the corona of volume 1 cm^3 at the point P (Fig. 59), which receives radiation $I_\odot$ dω from an element dS of the photosphere, scatters into unit solid angle in the direction of the observer an amount of energy

$$\frac{e^4}{2\,m^2\,c^4}\,n_e(r)\,I_\odot\,(1+\cos^2\theta)\,\frac{\mathrm{d}\omega}{4\,\pi}\,. \tag{22.10}$$

The total amount of energy scattered by unit volume of the corona into unit solid angle in the direction of the observer is obtained by integration:

$$j(r) = \frac{e^4}{2\,m^2\,c^4}\,n_e(r)\,R_\odot\int I_\odot\,(1+\cos^2\theta)\,\frac{\mathrm{d}\omega}{4\,\pi}\,, \tag{22.11}$$

where the integration is over the solid angle subtended by the Sun at the given point in the corona. This equation is similar to (22.6). Using (22.4), and solving complicated integral equations, these authors have found the values of $n_e(r)$ given in the second column of Table 15. These results differ from those of VAN DE HULST, n_e being up to 30% greater. However, these values also are a very rough approximation, since in reality the density of the corona varies with time and with position angle.

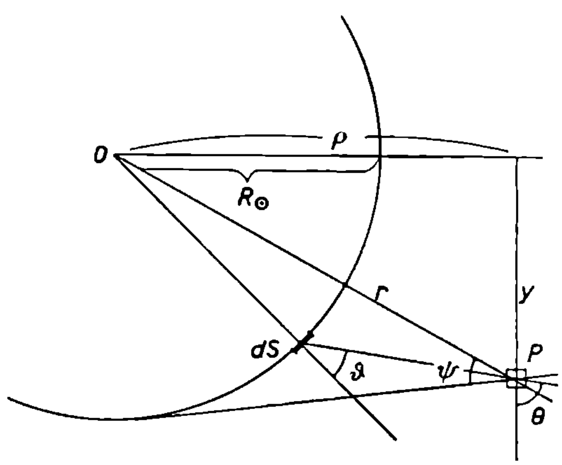

FIG. 59

Table 15

r	Corona with spherically symmetric density distribution	Maximum corona	Minimum corona equatorial	Minimum corona polar
1·00	519 × 10^6	403 × 10^6	227 × 10^6	174 × 10^6
1·03		316	178	127
1·06		235	132	87·2
1·1	156	160	90·0	53·2
1·2	74	70·8	39·8	16·3
1·3	41·4	37·6	21·2	5·98
1·5	18·3	14·8	8·3	1·41
1·7	8·8	7·11	4·0	0·542
2·0	3·36	2·81	1·58	0·196
2·6	0·73	0·665	0·374	0·040
3·0		0·313	0·176	0·017
4·0		0·090	0·050	0·004

VAN DE HULST, using the work of V. B. NIKONOV on the variation of the total brightness of the corona, has attempted to avoid such approximations. He considers separately the corona at maximum and minimum, taking the ratio of their total brightnesses as 1·84. He assumes the maximum corona to be spherically symmetrical, and the minimum corona to consist of an equatorial part, occupying 0·7 of the limb, and polar parts.

After a critical discussion of BAUMBACH's original data, and using the ratio of the brightnesses of the equatorial and polar parts of the corona obtained at the 1923 eclipse, VAN DE HULST derived from these data the brightness of the maximum corona and of the equatorial part of the minimum corona. The Fraunhofer component of the radiation was excluded, and from measurements of photographs of the 1900 eclipse he obtained the law of decrease of the density with ϱ for the polar region.

From all these results VAN DE HULST calculated $n_e(r)$ for the maximum corona and for the equatorial and polar parts of the minimum corona. The dependence of s_e on θ was taken into account in the calculations. The results are given in the third, fourth and fifth columns of Table 15.

The systematic difference between the results of A. F. BOGORODSKIĬ and N. A. KHINKULOVA and those of VAN DE HULST is apparently explained by the fact that VAN DE HULST, being unaware of the latest observations by the NIKONOVS, which gave a very low value for the total brightness of the minimum corona in 1945, took too small a value for the ratio of the brightnesses of the maximum and minimum corona. Moreover, VAN DE HULST did not use the observations of 1940, when the brightness of the corona was much greater than the mean brightness of the minimum corona which he took. Besides n_e, VAN DE HULST calculated the degree of polarisation of the corona at maximum and minimum, as a function of ϱ. A comparison with observation shows an agreement in the form of the function; quantitative discrepancies are apparently explained by the reasons mentioned above.

The anomalous extent of the corona was for long unexplained. If the corona were in hydrostatic equilibrium at the boundary temperature of the Sun, its extent would be some hundreds of kilometres. In 1947 H. ALFVÉN, assuming that the corona is in fact in hydrostatic equilibrium (as was later confirmed by a number of investigations), calculated the temperature T_k from the observed density distribution, assuming the corona to consist of ionised hydrogen. He obtained for T_k a value of about 1,000,000°, the temperature falling slowly with distance in the outer corona. His considerations are briefly as follows.

The condition of hydrostatic equilibrium, without taking account of radiation pressure, can be written in the form

$$\mathrm{d}p/\mathrm{d}r = -g\,\varrho\,. \tag{22.12}$$

Since the corona, on the above hypothesis, consists only of hydrogen atoms, and these completely ionised, the total number of particles in 1 cm^3 is $2\,n_e$. Hence

$$p = 2\,n_e\,kT\,, \quad \varrho = n_e\,m_{\mathrm{H}}\,. \tag{22.13}$$

From these formulae it is not difficult to determine the temperature from the observed density of the corona. The result is not significantly changed when the presence of other elements (helium, etc.) in the corona is taken into account.

3. **The identification of coronal lines.** For a long time the emission lines of the inner corona were ascribed to a hypothetical element called coronium. However, as Mendeleev's Table was filled up, attempts were repeatedly made to identify these lines with lines of elements already known. In 1939 W. Grotrian found that the wave numbers of forbidden transitions between sub-levels of the ground states of Fe X and Fe XI were close to the wave numbers of two coronal lines. Later, B. Edlén carried out a large amount of work on the identification of other coronal lines. He obtained the spectra of highly ionised elements, lying in the far ultra-violet region, in a spark discharge in a vacuum spectrograph. Having determined the wavelengths of the permitted lines and constructed from them a term diagram, he determined from this the wavelengths of the forbidden transitions between close levels. By this means two more lines were identified with transitions of Ca XII and Ca XIII. Edlén did not succeed in obtaining experimentally the lines of more highly ionised elements, and resorted to a theoretical extrapolation of the splitting of multiplets in various isoelectronic series (i. e. series of ions with increasing atomic number and degree of ionisation, having the same numbers of electrons). By this means the term differences of highly ionised atoms were determined and compared with the observed wavelengths.

In this way 19 lines out of 24 were identified with lines of highly ionised ions (the line 5694·4 was incorrectly identified, as has been shown by I. S. Shklovskiĭ; in the table it is listed as unidentified). A complete list of coronal lines and their identifications is given in Table 16.

Table 16

λ	Intensity: Grotrian	Intensity: Shaĭn	Intensity: Lyot	Identification	A sec^{-1}	$\varepsilon_{r,k}$	χ_r
3328	1·0			Ca XII $2s^2\,2p^5\,{}^2P_{\frac{1}{2}}—{}^2P_{\frac{3}{2}}$	488	3·72	589
3388·1	16			Fe XIII $3s^2\,3p^2\,{}^1D_2—{}^3P_2$	87	5·96	325
3454·1	2·3						
3601·0	2·1			Ni XVI $3s^2\,3p\,{}^2P_{\frac{3}{2}}—{}^2P_{\frac{1}{2}}$	193	3·44	455
3642·9				Ni XIII $3s^2\,3p^4\,{}^1D_2—{}^3P_1$	18	5·82	350
3800·8							
3986·9	0·7	4·7		Fe XI $3s^2\,3p^4\,{}^1D_2—{}^3P_1$	9·5	4·68	261
4086·3	1·0	1·0		Ca XIII $2s^2\,2p^4\,{}^3P_1—{}^3P_2$	319	3·03	655
4231·4	2·6	3·5		Ni XII $3s^2\,3p^5\,{}^2P_{\frac{1}{2}}—{}^2P_{\frac{3}{2}}$	237	2·93	318
4311							
4359				A XIV $2s^2\,2p\,{}^2P_{\frac{3}{2}}—{}^2P_{\frac{1}{2}}$	108	2·84	682
4567	1·1						
5116·0	4·3		2·2	Ni XIII $3s^2\,3p^4\,{}^3P_1—{}^3P_2$	157	2·42	350
5302·9	100	100	100	Fe XIV $3s^2\,3p\,{}^2P_{\frac{3}{2}}—{}^2P_{\frac{1}{2}}$	60	2·34	355
5446							
5536				A X $2s^2\,2p^5\,{}^2P_{\frac{1}{2}}—{}^2P_{\frac{3}{2}}$	106	2·24	421
5694·4			1·2				
6374·5	8·1	40·9	18	Fe X $3s^2\,3p^5\,{}^2P_{\frac{1}{2}}—{}^2P_{\frac{3}{2}}$	69	1·94	233
6701·8	5·4	7·7	2·0	Ni XV $3s^2\,3p^2\,{}^3P_1—{}^3P_0$	57	1·85	422
7059·6			2·2	Fe XV $3s\,3p\,{}^3P_2—{}^3P_1$		31·7	390
7891·9			13	Fe XI $3s^2\,3p^4\,{}^3P_1—{}^3P_2$	44	1·57	261
8024·2			0·5	Ni XV $3s^2\,3p^2\,{}^3P_2—{}^3P_1$	22	3·39	422
10,746·8			55	Fe XIII $3s^2\,3p^2\,{}^3P_1—{}^3P_0$	14	1·15	325
10,798·0			35	Fe XIII $3s^2\,3p^2\,{}^3P_2—{}^3P_1$	9·7	2·30	325

The second, third and fourth columns of the table give the results of measurements of the relative intensities of the lines during eclipse (Grotrian, G. A. Shaĭn [142]) and outside eclipse (B. Lyot). The sixth column gives the transition probabilities, the seventh gives the excitation potential of the initial level (the line at 7059·6 Å is formed by transitions between sub-levels of the second state), and the eighth gives the ionisation potential relative to the preceding ionisation state.

The following circumstance may serve as confirmation of the correctness of Edlén's identification. Lyot divided the observed lines into groups such that lines in the same group exhibit a similar distribution along the limb of the Sun at a given moment. Independently, G. A. Shaĭn

[141] divided the lines into groups according to their simultaneous strengthening or weakening in intensity during various eclipses. Despite the fact that the principle of the division into groups was completely different, the composition of the groups was approximately the same. This means that the intensities of the lines depend on the physical conditions and not on the chemical composition, so that the lines of a single group, under the same physical conditions, behave in the same way. A comparison of these groups with Table 16 shows that the ions of each group have nearly equal ionisation potentials, on which the degree of ionisation of the atoms and the behaviour of the lines depend.

Why can forbidden lines be observed in the corona? Their appearance requires, firstly, a low density of matter (in order that the time interval between two successive collisions of the second kind for one ion should be greater than the lifetime of an ion in the corresponding metastable state) and, secondly, a low density of radiation which transfers the atoms upwards from the metastable level. The first condition is fulfilled in the corona, since the probabilities of forbidden transitions are relatively large. The second condition is fulfilled because the excitations of an ion from a metastable level to higher non-metastable states require a large amount of energy, of the order of tens of electron-volts, and the intensity of solar radiation is small in the far ultra-violet region.

Wherein lies the cause of such a high degree of ionisation?

4. **The ionisation of atoms in the corona.** It may be regarded as now firmly established that the extremely high ionisation of the atoms in the corona is explained by the high kinetic temperature, which reaches a million degrees. The absence of absorption lines in the spectrum of the inner corona, and the small density gradient, also indicate a high temperature in the corona.

The emission lines in the corona have a half-width of about 0·5 Å. Since the lines are forbidden ($\Delta l = 0$), their natural width δ_{ik} is extremely small and need not be taken into account. The random (apparently turbulent) velocities of motion of the matter in the corona do not exceed 5 to 10 km/sec and again cannot explain the observed half-widths of the lines. Hence these can be accounted for only by thermal motions, and they correspond to a temperature of over 1,000,000°.

Finally, observations of the radio emission of the Sun, which we shall discuss at the end of this chapter, also indicate a high kinetic temperature of the corona.

The vast difference between the temperature of solar radiation in the corona and its kinetic temperature explains the marked deviation of the corona from the state of thermodynamic equilibrium and the unusual physical processes in it. The distribution of the atoms among the ionisation and excitation states has to be obtained not from thermodynamic

formulae, but from the condition for a steady state as regards the corresponding elementary processes (the numbers of any two converse processes are equal).

For hydrogen, which has only one ionisation state, the condition for a steady state has the form

$$Z_{\mathrm{pi}} + Z_{\mathrm{ci}} = Z_{\mathrm{pr}} + Z_{\mathrm{cr}}\ , \tag{22.14}$$

where the values of the first three terms are given by equations (8.9), (8.18) and (8.26). The number of triple collisions (Z_{cr}) is small because of the low density of the corona, and can be neglected. A comparison of the two terms on the left-hand side [their ratio is given by (8.23)] shows that, under the conditions in the corona, $Z_{\mathrm{ci}} \gg Z_{\mathrm{pi}}$ if χ_r exceeds 7 or 8 eV.

Thus, in the corona the *ionisation by collisions* is balanced by the *recombinations with emission of a quantum*; these processes are not converse ones, and consequently the ionisation formula here cannot be Saha's formula. In particular, the concentration of free electrons n_e does not appear in this formula, since both the number of ionisations by electron collisions and the number of recombinations are proportional to n_e. Moreover, the ionisation formula will be different for different atoms, since the nature and magnitude of the effective cross-sections $q_{r,k}(v)$ for ionisation by collisions depend on the particular structure of the atom. The ionisation of hydrogen and of metals in the corona has been considered by I. S. SHKLOVSKIĬ.

Let us consider the ionisation of hydrogen. The cross-section for ionisation of a hydrogen atom in the ground state by a fast electron (whose energy considerably exceeds the ionisation potential) is known from quantum mechanics, and is

$$q_0(v) = \frac{2\,\pi\, e^4}{m_e\, v^2\, \chi_0}\, 0{\cdot}285\, \log_e \left(\frac{2\, m_e\, v^2}{0.04\, \chi_0}\right)\ . \tag{22.15}$$

Substituting $q_0(v)$ in (8.18), integrating by parts and replacing $\frac{1}{2}\, m_e\, v_k^2$ by χ_0, we obtain

$$Z_{\mathrm{ci}} = n_e\, n_0\, 2 \sqrt{\left(\frac{2\,\pi\, m_e}{kT_e}\right)}\, 0{\cdot}285\, \frac{e^4}{m_e\, \chi_0} \left[e^{-\chi_0/kT_e} \log_e 100 + E_1\left(\frac{\chi_0}{kT_e}\right)\right]\ , \tag{22.16}$$

where the exponential integral $E_1(\chi_0/kT_e)$ is defined by (6.35); we do not take into account the ionisation from excited levels, since the excitation of hydrogen is small under the conditions in the corona. The value of Z_{pr} is determined by equation (8.26), where β_ν, by (5.55) and (5.24), is

$$(\beta_\nu)_k = \frac{g_{0,k}\, \nu^2\, h^2}{u_1\, c^2\, m_e^2\, v^2}\, \frac{32\,\pi^2\, e^6\, R}{3\sqrt{3}\, c h^3\, \nu^3\, k^5}\, g'\ . \tag{22.17}$$

Here $g_{0,k} = 2\,k^2$, ν is the frequency emitted in the recombination, which is related to the velocity v of the electron by (5.8), $u_1 = 1$ for ionised hydrogen, and $R = \nu_1$ (the frequency of the Lyman series limit) is given by the expression (5.25).

Substituting all these values in (8.26), neglecting stimulated recombinations ($c^3\,\varrho_\nu/8\,\pi\,h\,\nu^3 \ll 1$) and using the equation $\chi_{0,n} = \chi_0/n^2$, we obtain

$$Z_{\rm pr} = n_e\,n_1\,K\,T_\varepsilon^{-3/2} \sum_{n=1}^{\infty} \frac{1}{n^3} \int\limits_{v=0}^{\infty} \frac{e^{-m_e v^2/2\,kT_\varepsilon}}{(\chi_0/n^2\,kT_\varepsilon) + (m_e\,v^2/2\,kT_\varepsilon)}\, \mathrm{d}\left(\frac{\chi_0}{n^2\,kT_\varepsilon} + \frac{m_e\,v^2}{2\,kT_\varepsilon}\right)$$

$$= n_e\,n_1\,K\,T_\varepsilon^{-3/2} \sum_{n=1}^{\infty} e^{\chi_0/n^2 kT_\varepsilon}\, n^{-3}\, E_1\,(\chi_0/n^2\,k\,T_\varepsilon)\,, \tag{22.18}$$

where

$$K = \frac{2^4}{3\sqrt{3}\,\pi}\,\frac{8\,\pi^2\,e^2\,R^2}{m_e\,c^3}\,\frac{h^3}{(2\,\pi\,m_e\,k)^{3/2}} = 3{\cdot}2 \times 10^{-6}\,. \tag{22.19}$$

For the higher terms of the series, stimulated recombinations cannot be neglected, but the value of the sum is largely determined by the first term.

Substituting (22.16) and (22.18) in (22.14), and using (5.25) and the equation $\chi_0 = h\,R$, we find the ionisation formula for hydrogen

$$\frac{n_1}{n_0} = \frac{3\sqrt{3}\,.\,0{\cdot}285}{2^7\,\pi^5}\,\frac{c^3\,h^5\,kT_\varepsilon}{m e^{10}}\,\frac{e^{-\chi_0/kT_\varepsilon}\log_e 100 + E_1\,(\chi_0/kT_\varepsilon)}{\sum\limits_{n=1}^{\infty} e^{\chi_0/n^2\,kT_\varepsilon}\,n^{-3}\,E_1\,(\chi_0/n^2\,kT_\varepsilon)}\,. \tag{22.20}$$

As we have already said, the ionisation is independent of n_e. For $T_\varepsilon = 625{,}000°$, $n_1/n_0 \approx 10^6$; for $T_\varepsilon = 1{,}560{,}000°$, $n_1/n_0 \approx 3 \times 10^6$. If the ionisation were determined by Saha's formula with $T = T_\varepsilon$, n_1/n_0 would be almost 10^6 times greater, and for $T = 5000°$ $n_1/n_0 = 0{\cdot}07$.

For an atom with many electrons, such as iron, the conditions of ionisation equilibrium take the form of a system of equations, each of which expresses the equality of the numbers of elementary processes which lead to the appearance and disappearance of atoms in a given state of ionisation. In these equations, as in (22.14), triple collisions can be neglected, as can photo-ionisation (for all ionisation states, besides the first) and ionisation from excited levels by collisions.

The effective cross-sections for ionisation by electron collisions appear in the equations. The problem is complicated by the fact that, for highly ionised atoms, the ionisation energy is greater than the mean thermal energy of an electron, so that the ionisation is effected mainly by "slow"

electrons, and the exact values of the effective cross-section are unknown in this case. We shall use the formula (derived from classical considerations)

$$q_r(v) = \frac{\pi e^4}{\frac{1}{2} m_e v^2} \frac{1}{\frac{1}{2} m_e v_1^2 + \chi_r}, \tag{22.21}$$

where v is the velocity of the colliding electron and v_1 that of the electron detached from the atom. In our case $\frac{1}{2} m_e v_1^2 \ll \chi_r$. This formula gives about the same value for $q_r(v)$ as does the approximate quantum-mechanical formula, and roughly agrees with the experimental results for helium.

For the number of recombinations, we take the expression

$$Z_{\mathrm{pr}} = n_e n_{r+1} K T_\varepsilon^{-3/2} Z^4 \sum_{n=n_1}^{\infty} e^{Z^2 h R/n^2 k T_\varepsilon} n^{-3} E_1(Z^2 h R/n^2 kT_\varepsilon), \tag{22.22}$$

which is a generalisation of (22.18) to the case of a hydrogen-like ion with nuclear charge Ze. In reality, the ions we are considering are not hydrogen-like. In this case the quantity Ze can be regarded as some effective charge acting on the outer electron. In general, the application of (22.22) to a non-hydrogen-like ion may lead to considerable errors. However, the majority of the recombinations take place to excited levels. These are, in practice, hydrogen-like. A numerical solution of the system of equations of the steady state for Fe and Ni shows that to each value of T_ε there corresponds a definite ionisation state which is most often found.

At $T_\varepsilon = 600{,}000°$, iron is chiefly in the state Fe X, and at $T_\varepsilon = 1{,}200{,}000°$ in the state Fe XIV*. Thus the "red" and "green" coronal regions correspond to the temperatures mentioned. The fact that observation sometimes shows an intensification of both the red and the green lines at the same point in the corona is explained by the fact that the line of sight passes successively through green and red regions. Observation shows that the regions of radiation from Ni XIII and Ni XV largely coincide with those of radiation from Fe XIII and Fe XIV respectively. This is explained by the fact that the numbers of the respective ions in the given ionisation state reach a maximum at the same temperatures.

5. **The chemical composition of the corona.** The corona is transparent to its own monochromatic radiation, because the absorption coefficient for the forbidden lines is small, and the total optical thickness of the corona in these lines is many orders of magnitude less than unity. Thus, from the intensities of the emission lines we can form an estimate of the

* A. Werner has shown that a new value of the recombination coefficient for Fe XIV gives $T_\varepsilon \sim 2{,}000{,}000°$ for the green regions.

total number of atoms which emit in a given line. Calculations relating to the chemical composition of the corona were first performed by I. S. SHKLOVSKIĬ.

Let $I_\lambda(\varrho)$ be the energy emitted by 1 cm^2 of the projected corona in unit solid angle, integrated over the whole line. Observation usually gives the equivalent widths of the emission lines, referred to the neighbouring continuous spectrum of the same part of the corona, $\Delta\lambda(\varrho)$, i. e. the interval of the adjoining continuous spectrum which contains the same amount of energy as does the line. Knowing the intensity distribution $I(\varrho)$ in the continuous spectrum, determined by (22.3) (here it is unnecessary to exclude the Fraunhofer component of the radiation, since $\Delta\lambda$ refers to the whole continuous spectrum), we find $I_\lambda(\varrho) = \Delta\lambda(\varrho)\; I(\varrho)$. An integral equation similar to (22.4) enables us to calculate the emission coefficient $j_\lambda(r)$, which in turn determines the number of atoms $n_k(r)$ in the initial state:

$$4\,\pi\, j_\lambda(r) = n_k(r)\;(A_{ki} + \varrho_{\nu_{ik}}\, B_{ki})\, h\, \nu_{ik}\,. \tag{22.23}$$

The values of $n_k(r)$ have been calculated from all the existing spectrophotometric observations of the monochromatic radiation of the corona. As a result it is found that, for $r = 1.2$, $n_k(r)$ has values from some tenths to some tens of atoms per cm^3 for various ions.

Next, let us pass from the number of excited atoms to the total number of atoms in the ionisation state considered. To do this, we must examine various mechanisms of excitation. First of all, we must take account of excitation by electron collision. The number Z_{ce} of such excitations is determined by a formula like (8.18), where $q_{r,k}(v)$ denotes the effective cross-section for excitation. The calculation of such cross-sections for small relative velocities is an extremely difficult problem. We shall use the formula for the excitation of sub-levels of the same configuration of O III, calculated by D. H. MENZEL and M. H. HEBB on the assumption of Russell-Saunders coupling*. Since excitation of sub-levels of the ground state also takes place in the corona, we can assume, as a first approximation, that the chief property of O III (the large value of $q_{r,k}(v)$, which is of the order of ten times the value found from the kinetic theory of gases) holds for the coronal ions also. A comparison of theory and observation allows us to test this assumption. The population of an excited level depends both on excitations by collisions and on recombinations followed by cascade transitions to the level in question. It is easy to see that $Z_{ce} > Z_{ci}$, since $(q_{r,k})_e > (q_r)_i$, and hence recombinations to excited levels, which are fewer in number than Z_{ci}, need not be considered; they are rare in comparison with direct excitations.

* M. J. SEATON has recently obtained new values for the excitation cross-sections, and these are somewhat less than MENZEL's values.

In the corona (unlike nebulae) excitations taking place owing to the absorption of radiation from the photosphere cannot be neglected. Direct calculation (the quantities B_{ik} being known) shows that, for the lines $\lambda = 5303$ and $\lambda = 6374$ Å, photo-excitations will predominate for r greater than $2{\cdot}2\ R_\odot$ and $1{\cdot}6\ R_\odot$ respectively, where n_e is already considerably less than in the inner corona, but the flux of radiation is not so strongly diminished. In this case, the emission coefficient over the whole line $j_\lambda(r)$ in the inner corona will be proportional to $n_e\, n_{r,1}$, and in the outer corona to $n_{r,1}$. If the ionisation (determined by T_ε) and the relative content of iron do not vary with height, then $n_{r,1}$ is proportional to the number of hydrogen atoms in 1 cm³ and is consequently proportional to n_e. This means that in the inner corona $j_\lambda(r) \sim n_e^2$, and in the outer corona $j_\lambda(r) \sim n_e \sim j(r)$. Observations (by C. W. ALLEN) in fact show that, from $r = 2{\cdot}2\ R_\odot$ (approximately the value obtained from calculation) outwards, $j(r)$ for $\lambda = 5303$ Å is proportional to $j(r)$ for "white" light. This confirms the correctness of the value taken for $(q_{r,k})_e$.

Knowing the excitation mechanism, we can calculate the distribution of the atoms among the excitation states, and so find n_r from $n_{r,k}$. Having data for the lines of atoms in several ionisation states, and having calculated, from the theory of ionisation, the numbers of atoms in the "invisible" ionisation states (for the red and green regions separately), we can determine the chemical composition. The concentration of H is given by $n_e(r)$.

It has been found that the relative content of Fe and H is constant throughout the corona, and is close to the content of these elements in the reversing layer. This result is very important; it shows that the elements in the corona are mixed and that there is not diffusion equilibrium, in which the heavier atoms would fall to the lower layers. This mixing is apparently caused by slow turbulent motions in the corona. It is natural to suppose that the elements are mixed in the intermediate layers of the solar atmosphere (the chromosphere and the prominences) also.

6. The ultra-violet radiation of the corona and the chromosphere [149]. The radiation of the corona does not terminate at the limits of the visible spectrum, but extends on both sides of it. This is primarily radiation beyond the principal series limits of H, He, He II and other elements, in consequence of recombinations. It is given by formula (8.38).

Radiation will also be produced by free-free transitions. Its intensity depends only on the density and on the velocities of the particles, and is determined for hydrogen by the formula

$$4\,\pi\, j_\nu\, \mathrm{d}\nu = n_1\, n_e\, \frac{kK}{2\, R T_\varepsilon^{\frac{1}{2}}}\, g''\, e^{-h\nu/kT_\varepsilon}\, \mathrm{d}\nu\,. \qquad (22.24)$$

[The expression (22.24), for the particular case of thermodynamic equilibrium, can be obtained from (5.65) and (3.10).] The second source of emission (free-free transitions) for hydrogen under the conditions in the corona, i. e. at a very high kinetic temperature, is about three times as efficient as the first. On account of the high degree of ionisation, the corona is transparent even to radiation beyond the Lyman series limit. Its radiation therefore emerges unimpeded.

Thus the intensity of the continuous radiation of the corona is proportional to $P = \int\limits_0^\infty n_1 n_e \, \mathrm{d}y = \int\limits_0^\infty n_e^2 \, \mathrm{d}y$ (see Fig. 59). At the limb, where the line of sight is a tangent to the surface, P is greatest and the brightness is a maximum. The structure of the corona (the presence of jets and rays, the non-uniform density along the limb) increases P for a given mean electron density, and consequently increases the brightness. At the distance of the Earth, the total flux H_1 of ultra-violet radiation from hydrogen in the range from ν_1 to $\nu = \infty$ is 4×10^{-2} erg/cm^2 sec. Similar calculations have shown that the amount of radiation due to helium is 1·4 times as great as that due to hydrogen. The part played by radiation due to the remaining elements is negligible.

Apart from the continuous spectrum, the corona also has a line spectrum in the ultra-violet region. This arises from permitted transitions between excited levels and the ground level of Fe and Ni ions in electron collisions. Each excitation by an electron collision leads to the emission of a quantum; the number of these is determined by a formula like (8.18), so that j_λ for the monochromatic radiation also is proportional to $n_r \, n_e \sim n_e^2$. The cross-section for excitation by fast electrons (the thermal energy of an electron is about 120 eV) is

$$(q_{r,k})_e = \frac{2 \pi e^4}{m v^2 \varepsilon_{r,k}} f_{0,k} \log_e \left(\frac{2 m v^2}{\varepsilon_{r,k}} \right), \tag{22.25}$$

where $f_{0,k}$ is the oscillator strength for the transition (close to 1 for resonance lines).

We shall give the results for a number of elements which are fairly common in the Sun. For instance, Ne VIII (neon reaches this state in "red" regions) and Mg X ("green" regions) have resonance doublet transitions 776 to 768 and 625 to 610 Å respectively. Since either the red or the green line is observed in the greater part of the corona, we can calculate the total emission of energy in the two pairs of ultra-violet lines by integrating the quantity $4 \pi j_\lambda = h \nu Z_{ce}$ over the whole volume of the corona. Thus we suppose that either the Ne VIII line or the Mg X line is emitted at every point of the corona.

Having given the relative content of Mg and Ne, we find for the total quantity of monochromatic radiation formed in the entire corona the value $E = 6 \times 10^{27}$ erg/sec. If half the radiation returns to the Sun, the flux of hard monochromatic radiation at the distance R_1 of the Earth is

$$H_2 = \tfrac{1}{2}\, E/4\,\pi\, R_1^2 \approx 1 \text{ erg/cm}^2 \text{ sec.} \tag{22.26}$$

The value of H_2 is large enough to explain the observed ionisation of the upper layers of the Earth's atmosphere, although it is possible that the numerical value of H_2 in (22.26) is too large. We have already mentioned the ionisation of the chromosphere and prominences by the radiation of the corona.

The upper layers of the chromosphere, where hydrogen is almost completely ionised, must also emit ultra-violet radiation which is due to recombinations to the ground level. Free-free transitions do not play an important part at the low kinetic temperature of the chromosphere.

If the ionisation of the chromosphere is effected only by the radiation of the corona, the chromosphere plays, as it were, the part of a reflector, which reflects part of the downward radiation; the flux from it is thus less than from the corona. If the temperature of the various regions of the upper chromosphere (above 7000 to 8000 km) exceeds 20,000°, and thus determines the ionisation, the radiation from these regions will also be sufficient to ionise the Earth's atmosphere.

The energy of the ultra-violet radiation, and of the greater part of the visible monochromatic radiation of the corona, is derived from the thermal energy of the electron gas. In nebulae this loss is made up from the radiation of the central star. In the corona, this possibility is excluded, because of the low temperature of the photosphere, and another source must be sought. I. S. Shklovskiĭ has advanced the hypothesis that the source of heat is formed by currents which arise under the action of electric fields (for instance, an induction field). The heat q evolved by these is determined by formula (18.37). This energy must balance, on the one hand, the flux of energy of radiation $4\,\pi\, j_\lambda = n_e^2\, \Psi(T_\varepsilon)$ and, on the other hand, the flux of energy due to heat conduction in the lower layers of the atmosphere and in the outer layers of the corona.

The amount of heat evolved is

$$\dot{q} = \lambda\, E^2 = n_e^2\, \Psi(T_\varepsilon) - \gamma\, \mathrm{d}^2 T_\varepsilon/\mathrm{d}r^2\,, \tag{22.27}$$

where γ is the thermal conductivity of the electron gas, and $\Psi(T_\varepsilon)$ is some function of T_ε. From formula (22.27) it may be calculated that, to maintain the temperature of the corona at its observed value, a field of intensity $E \approx 10^{-8}$ volts/cm is required, the two terms on the right-

hand side being then of the same order of magnitude. This small value of the field intensity needed to supply the heat is explained by the small heat loss of the corona, which is due to the small value of n_e and the small temperature gradient. G. A. LEĬKIN has shown that the field E may be due to the rotation of the corona in the interstellar magnetic field $H \approx 10^{-5}$.

In the chromosphere n_e is a thousand times greater, and hence even stronger fields have a relatively slight heating effect. The question arises why the temperature of the corona is approximately constant although the fields may vary by a factor of a hundred or even a thousand. Furthermore, it is not clear why T_ε decreases towards the edges of the corona, although n_e, and consequently the heat loss, is less there. As has been shown by S. B. PIKEL'NER [124], this is explained by the surface effect of cooling of the corona by the outflow from it (*dissipation*) of the fastest particles, which have velocities greater than the parabolic. The outflow of particles takes place mainly from the layer $1{\cdot}75 < r < 2{\cdot}75$; outflow from deeper layers is prevented by the resistance of the medium, and in the higher layers there are few collisions which "create" fast particles. The dissipation is proportional to the fraction of particles which are fast, i. e. to $\exp\,[-m\, v_\infty^2/2\, kT_\varepsilon]$, where v_∞ is the velocity necessary to overcome resistance and gravitation (the parabolic velocity). As T_ε increases, this fraction rapidly increases. The dissipation is, as it were, a temperature control. If at some point in the corona the temperature is raised above 1,500,000°, the thermal conductivity transmits the part $\gamma(\mathrm{d}^2T_\varepsilon/\mathrm{d}r^2)$ of the energy to the higher layers, and so this part of the energy is lost into space by dissipation. V. A. KRAT [71] first carried out a calculation of the dissipation of particles from the corona, taking into account its high kinetic temperature.

Thus it is not accidental that the thermal velocities of protons in the corona are comparable with the parabolic velocity (as is easily shown by calculation). This circumstance also explains the great extent of the corona, which is due to the closeness of the thermal and parabolic velocities of the protons. In other stars, the same heating mechanisms must undoubtedly exist, since the Sun is an "ordinary" star. The upper limit to the temperature of the corona is determined by the ratio of the thermal and parabolic velocities (i. e., by the mass and radius of the star). For example, in the corona of a star of class gK 5, $T_\varepsilon \approx 160{,}000°$; in a star of class cB 3, $T_\varepsilon \approx 2{\cdot}7 \times 10^6$. The density distribution in the corona of a star will be similar to that in the solar corona if we take the radius of the star as the unit in measuring the extent of the corona; thus a giant star must have a very extended corona. The ultra-violet radiation of such a corona may cause the observed anomalous excitation of atoms with high excitation potentials, such as hydrogen atoms.

7. **The radio emission of the Sun.** Until recently, all observations of the radiation of celestial objects were made in frequencies in the neighbourhood of the visible region of the spectrum, from $\lambda = 2900$ Å to the comparatively far infra-red. This range is limited on the short-wave side by the absorption in the ozone layer of the Earth's atmosphere, and on the long-wave side by the water-vapour absorption bands. Moreover, the sensitivity of radiation receivers is very small for $\lambda > 30{,}000$ Å. However, there is another transparent "window" for wavelengths of from 1 cm to 12 m, the **ultra-short radio waves.**

The invention of sufficiently sensitive and directional radar apparatus has made it possible to detect the radiation of the Galaxy and of the Sun in this wavelength range. The intensity of the radiation can be measured in terms of an arbitrary quantity, the **radio temperature** T_r, which is the temperature of a black body of the same angular dimensions which emits the same flux in the given frequency. The radio temperature of the Sun at wavelengths from 1 to 10 cm increases from 10,000° to 20,000° as the wavelength increases, and is almost constant with time. At wavelengths of about 50 cm, the radio temperature of the Sun is measured in hundreds of thousands of degrees. At a wavelength of 1 m, the radio temperature varies markedly with time, particularly in the years of maximum solar activity, occasionally reaching values of 10^{13} degrees. At these wavelengths T_r never falls below 10^6 degrees, and in years of minimum activity remains almost constant at about this value. It is convenient to divide the radiation at metre wavelengths into a **quiescent** or **thermal** and a **sporadic** component (which varies with time).

The thermal radiation has been investigated by I. S. Shklovskiĭ, and independently by V. L. Ginzburg and by D. F. Martyn. The absorption coefficient of radio waves in an ionised gas is determined by free-free transitions [see formula (5.65)] with $T = T_\varepsilon$. For the quantity $e^{-h\nu/kT}$ in formula (5.65) is practically equal to the ratio of the numbers of atoms in the upper and lower quantum states. The attenuation of a beam of rays in a given direction is determined by the difference between the energy actually absorbed (the coefficient B_{ik}) and the energy emitted by the atoms in the same direction, under the action of the quanta incident on them (the coefficient B_{ki}). Hence the decrease in the number of quanta when they traverse an element dh is given by the expression (which does not take into account the emission of the gas itself)

$$n_i \varrho_\nu B_{ik}\, dh - n_k \varrho_\nu B_{ki}\, dh = n_i \varrho_\nu B_{ik} \left(1 - \frac{n_k}{n_i} \frac{g_i}{g_k}\right) dh\,.$$

Using formula (5.9) for the subtrahend, we obtain $e^{-h\nu/kT}$. In the case considered, where the emission is accompanied by free-free transitions of the electron, the relative number of electrons with a given energy

is determined by Maxwell's formula with $T = T_\varepsilon$, and the ratio of the numbers of electrons with the given energy difference is $e^{-h\nu/kT_\varepsilon}$. The fact that we originally spoke of atoms in different quantum states, and later of the electron energy, should not cause any misunderstanding, since the emitting system in the case considered is an ion and an electron, and the energy of such a system is equal to that of the electron.

Since $h\nu/kT_\varepsilon \ll 1$, we can rewrite the absorption coefficient, assuming the corona to consist of hydrogen, in the form [see (5.65)]

$$\frac{\varkappa_\nu}{\varrho} = \frac{4\pi e^6}{3\sqrt{3}}\frac{1}{c}\frac{n_e^2}{(m\,kT_\varepsilon)^{3/2}}\frac{g''}{\nu^2} = \psi(\nu, T_\varepsilon)\, n_e^2\,, \tag{22.28}$$

where

$$\psi(\nu, T_\varepsilon) = 1{\cdot}65 \times 10^{-7}\left(\frac{10{,}000}{T_\varepsilon}\right)^{3/2}\frac{1}{\nu^2}\,.$$

It is easily calculated that the optical thickness τ_k of the chromosphere is greater than 10^3 even for $\lambda = 1$ cm. The corona has $\tau_k = 1{\cdot}6$ for $\lambda = 150$ cm. Hence the radiation which reaches us with $\lambda \geq 1$ m comes mainly from the corona.

The intensity of the emergent radiation is determined by formula (3.32). As a first approximation, it may be assumed that T_ε in the corona is constant and equal to T_k, while T_ε in the upper chromosphere, at the level where $\tau_\nu \approx 1$, is equal to some T_c. Then formula (3.32) can be rewritten in the form

$$I_\nu(\theta, 0) = B_\nu(T_c)\, e^{-\tau_k} + B_\nu(T_k)\,\left(1 - e^{-\tau_k}\right)\,, \tag{22.29}$$

where τ_k is measured along the ray path. For centimetre wavelengths, $\tau_k \ll 1$, the second term is small compared with the first, the disc is uniformly bright, and $T_r = T_c$. For wavelengths in the metre range, the first term is zero, and the brightness of the disc diminishes slowly towards the limb, the diameter of the disc being greater than that of the Sun. For decimetre wavelengths, the two terms are comparable near the centre of the disc, while at the limb $\tau_k \approx 1$ and the radiation resembles that of a black body with $T = T_k$. The Sun thus appears as a bright ring.

The increase of the radio temperature with wavelength is, of course, explained by the continually increasing part played by the emission of the corona.

Radio waves have two properties which are unlike those of ordinary radiation. The first is that their refraction coefficient in an ionised gas

$$n = \sqrt{\left(1 - \frac{e^2 n_e}{\pi m_e \nu^2}\right)} \tag{22.30}$$

is less than unity (the group velocity u of the waves is cn), while for metre wavelengths in the corona n may become zero, so that we have total reflection. Radiation can emerge only from layers lying above the level where $n = 0$. The path of a ray which does not pass normally to the surface is curved so that it is convex towards the Sun. For spherically symmetric layers, the equation of the path is the usual equation of refraction

$$rn \sin i = \text{constant}\,, \tag{22.31}$$

where i is the angle between the ray and the radius through the point. For this reason, τ_k must be measured not along a straight line, but along the path of the ray, and therefore equations (22.29) and (22.31) must be solved simultaneously. For centimetre wavelengths, the refraction in the corona is insignificant.

Calculations of the brightness distribution of the radio emission over the disc and of the radio temperatures of the Sun for various wavelengths have been carried out independently by several investigators.

A simultaneous measurement of T_r for several wavelengths allows us to obtain a system of equations (22.29) with the unknowns T_c and T_k (τ_k is calculated theoretically) and so to determine these latter from observation. Observations of T_r at centimetre wavelengths are one of the most reliable methods of determining the temperature of the upper layers of the chromosphere.

The second peculiarity of radio waves is the possibility of their polarisation as a result of the occurrence of double refraction in a plasma, in the presence of a magnetic field. The oscillations which take place in planes parallel and perpendicular to the field have different values of n, and consequently different velocities of propagation and different absorptions. Under these conditions, elliptic polarisation can occur.

As we have already said, the sporadic radiation is often observed at metre wavelengths, particularly in years of maximum activity. Recent observations have made possible a division of the sporadic radiation into a component which varies comparatively slowly, comes from the regions above spots, and has a fairly definite radial direction, and bursts — sharp, rapidly varying oscillations in intensity.

The sporadic radiation above spots is often elliptically polarised, because of the magnetic fields of the spots.

The bursts are divided into *short* (lasting about 1 sec), weak (exceeding the intensity of the thermal radiation by a factor of three or four) bursts and *long* (lasting several minutes), infrequent, powerful bursts, during which T_r for the Sun reaches values of 10^9, and sometimes of even 10^{13}, degrees. Since this radiation is not in equilibrium, the significance of these values is purely formal.

The large bursts are explained by I. S. SHKLOVSKIĬ as due to the natural oscillations of the plasma, occurring under the action of a flux of rapidly moving particles. It is well known that oscillations of electrons with a frequency

$$\nu = \sqrt{(e^2 \, n_e/\pi \, m_e)} \tag{22.32}$$

may arise in the plasma; this frequency depends only on n_e. These oscillations must cause electromagnetic oscillations of the same frequency. Calculations show that the waves corresponding to the electron concentration in the corona have a length of several metres; the higher layers emit longer waves. Large bursts frequently occur some minutes after chromospheric flares. In such cases the bursts first appear at the shorter wavelengths, and then at progressively longer wavelengths. The phenomenon gives the impression that the agency responsible for the oscillations gradually penetrates into higher and higher layers of the corona. A magnetic storm is frequently observed some days after a burst. From the delay time of bursts in different wavelengths, we can calculate the velocity of propagation of the current, knowing the distribution of $n_e(r)$; the result is 500 to 1000 km/sec. This confirms the hypothesis that the bursts are caused by a flux of geoactive corpuscles related, in particular, to chromospheric flares. The concentration of particles necessary to excite oscillations of the observed magnitude is found to be 10^2 to 10^3 per cm^3, which is even less than the value obtained from ionosphere data.

V. L. GINZBURG explains the sporadic emission above spots as being due to the emission by relativistic electrons in the magnetic field of the spot. The presence of relativistic electrons is indicated by the fact that the Sun emits cosmic rays at the time of large chromospheric flares and strong radio emission.

PART IV.
PLANETARY NEBULAE

In this part we shall begin our consideration of objects with bright lines in their spectra. Among these are the planetary nebulae, the novae, stars of the Wolf-Rayet, P Cygni, and Be types, and so on. As will be seen below, in all these cases we are concerned with the ejection of matter from hot stars. This process leads to the formation of very extended and rarefied envelopes. It is in these envelopes that the bright lines arise as a result of the transformation of the high-frequency radiation of the star into radiation of lower frequencies and, in particular, into radiation in the visible part of the spectrum.

The fact that the radiation of the envelopes in the visible part of the spectrum is produced at the expense of radiation in the ultra-violet region of the star's spectrum indicates a marked deviation of the state of the envelopes from thermodynamic equilibrium.

This distinguishes the envelopes of stars with bright spectral lines from the atmospheres of ordinary stars, where the assumption that thermodynamic equilibrium exists is adequate as a first approximation to reality. In studying the objects mentioned above, we shall not be able to use Boltzmann's and Saha's formulae to calculate the numbers of atoms in the various states, or Planck's formula to calculate the intensity of radiation in different frequencies. These quantities must be determined in each individual case by considering the elementary processes which take place in the actual envelopes. We shall usually make the assumption that the envelopes are in a *steady state*, i. e. that the distribution of atoms among the states and the radiation field in the envelope do not vary in the course of time. Here, of course, we need to know the probabilities of the various elementary processes, i. e. the probabilities of photo-ionisation, recombination, collision, etc., which are calculated in theoretical physics.

In consequence of the extremely low density of matter and radiation in planetary nebulae, the physical processes which take place in them are relatively simple.

Chapter 23. The mechanism of the radiation of the nebulae. The temperatures of their nuclei

1. **Observational data.** A **planetary nebula** is a luminous gaseous mass of fairly regular shape with a star lying in its centre, called the **nucleus of the nebula.** In a telescope, planetary nebulae most often appear as *round* or *oval discs*, very similar to those of the planets (whence the name). A considerable number of nebulae are observed as *rings* surrounding the nucleus. Some nebulae have a very complex structure (for instance, two intersecting rings).

The angular diameters of planetary nebulae rarely exceed one minute of arc. Many (very small or very distant) planetary nebulae show no discs at all in the telescope, and it can be established only from the form of the spectrum that the object in question is a planetary nebula. In recent years more than a hundred such "stellar nebulae" have been discovered, and the total number of all planetary nebulae at present known amounts to about 350.

The planetary nebulae are so distant that for none of them has it been possible to make a reliable determination of the parallax by the trigonometrical method. The mean distances of the nebulae we observe, found from the radial velocities by means of the theory of galactic rotation, are of the order of some thousands of parsecs. Correspondingly, the mean linear diameters of the nebulae are found to be of the order of some tens of thousands of astronomical units, and their mean absolute magnitudes are about 0^m.

As a rule, the planetary nebulae are much brighter than their central stars. The differences in stellar magnitude between the nebulae and their nuclei are almost always negative and sometimes reach -7^m. A value of about $+3^m$ is obtained for the mean absolute photographic magnitude of the nucleus. The nuclei of the nebulae are very hot stars. Their spectra belong to the classes O and Wolf-Rayet. However, it must be remarked that in luminosity the nuclei of the nebulae are much inferior to average stars of these classes, whose absolute magnitudes are about -3^m.

The spectra of planetary nebulae consist of bright lines of various atoms and ions. First of all, we must mention the presence of very intense lines of the Balmer series of hydrogen. Besides these lines, the Balmer continuum is observed in emission, owing to captures of free electrons by protons. As well as lines of neutral helium, there are present, in the spectra of many nebulae, lines of ionised helium also (for example 4686 Å); this indicates a very high degree of ionisation and excitation in these nebulae.

However, the brightest lines in the spectra of planetary nebulae are what are called the *principal nebular lines* N_1 and N_2, with wavelengths of 5007 and 4959 Å respectively. Another pair of very intense lines, 3726 and 3729 Å, is observed in the ultra-violet region of the spectrum. These two doublets are also found in the spectra of diffuse nebulae and novae, but they are dominant in the spectra of planetary nebulae. Since these lines, as well as a number of others found in the spectra of nebulae, could not be reproduced in the laboratory, they were formerly ascribed to an element "nebulium" not known on the Earth. However, in 1928 I. S. BOWEN succeeded in showing that the "nebulium" lines are in reality forbidden lines of elements known to us. Their presence in the spectra of nebulae reveals the peculiar physical conditions which exist there, and which have not been attained in terrestrial laboratories.

According to BOWEN's interpretations, the green doublet in the spectra of nebulae (the N_1 and N_2 lines) is caused by forbidden transitions of doubly ionised oxygen atoms, while the ultra-violet doublet is due to forbidden transitions of singly ionised oxygen atoms. The other lines of "nebulium" are likewise forbidden lines of oxygen, nitrogen, and some other elements in various stages of ionisation. The correctness of this identification is confirmed not only by the exact agreement of the frequencies of the observed lines with those found theoretically from the disposition of the energy levels in the atoms, but also by numerous other facts. As an example, we may mention that, according to BOWEN, the N_1 and N_2 lines have a common upper level, and the ratio of their intensities should always be equal to 3. Such a ratio of the intensities of these lines is in fact observed in the spectra of nebulae.

The *shape* of the spectral lines in planetary nebulae is of great interest. If the slit of the spectrograph is placed along a diameter of the nebula, the spectral lines are found to be narrow at the edges and wider in the centre. In some cases the lines are split into two components in the centre. This form of the spectral lines is explained by the expansion of planetary nebulae. The violet component of a spectral line is then formed by the part of the nebula which is approaching us, and the red component by the part receding from us. It is evident that the greatest separation of the components (or the greatest width of the line, in the case where the components coalesce) corresponds to twice the velocity of expansion of the nebula. These velocities, determined for a number of nebulae, have been found to be of the order of 10 to 20 km/sec. The fact that planetary nebulae are expanding leads to the hypothesis that they were at some time ejected from their nuclei.

2. **The origin of the radiation of the nebulae.** Passing now to the interpretation of the observational data, we shall first of all answer the question of the origin of the radiation of planetary nebulae. It is

natural to suppose that the radiation of the nebulae is due to their central stars. However, in this case it is not a simple reflection of light, since the spectra of the nebulae are not at all similar to those of the stars. Nor can there be resonance scattering in the nebulae, since the energy radiated by the nebulae in the separate lines considerably exceeds the energy in the corresponding parts of the continuous spectra of the nuclei. It is therefore necessary to assume that the nebulae radiate in the visible region of the spectrum at the expense of energy emitted by the nuclei in other regions of the spectrum. Since the temperatures of the nuclei are very high (over 30,000°), and consequently they have very great intensities in the ultra-violet region, the hypothesis suggests itself that the nebulae absorb the ultra-violet radiation of the nuclei and re-emit it in lower frequencies. We shall now show that this assumption is entirely correct.

Let us first consider the properties of the radiation which passes from the nucleus to a given point in the nebula. Let the nucleus radiate like a black body of temperature T_*. We denote by ϱ_ν^* the density of radiation in thermodynamic equilibrium at temperature T_*. This density is determined by Planck's formula

$$\varrho_\nu^* = \frac{8\pi h \nu^3}{c^3} \frac{1}{e^{h\nu/kT_*} - 1} . \tag{23.1}$$

It is evident that the density of radiation at the surface of the star is $\frac{1}{2}\varrho_\nu^*$, and it decreases with increasing distance from the star, according to the law

$$\varrho_\nu = W \varrho_\nu^* , \tag{23.2}$$

where the quantity W, which we have already introduced in Chapter 8, is the dilution coefficient; it is equal to the ratio of the solid angle subtended at the point concerned by the central star to 4π. If the radius of the star is r_*, the dilution coefficient at a distance r from its centre is, by (8.3),

$$W = \tfrac{1}{2}\left[1 - \sqrt{\{1 - (r_*/r)^2\}}\right] . \tag{23.3}$$

When $r_*/r \ll 1$, we have instead of (23.3)

$$W = \tfrac{1}{4}(r_*/r)^2 . \tag{23.4}$$

In planetary nebulae, the ratio r_*/r is of the order of 10^{-7}. Consequently, the density of radiation in a planetary nebula is reduced by a factor of approximately 10^{14} compared with the density of radiation at the surface of the star.

It is very important that, despite such a great decrease in the density of stellar radiation in the nebula, the relative energy distribution in its spectrum remains unchanged. In order to emphasise that the decrease in the density of radiation does not affect its spectral composition, we shall calculate the temperature corresponding to a given density of radiation in the nebula. Denoting this temperature by T_1, we have the integrated density of radiation $\varrho = a\, T_1^4$, where a is Stefan's constant. On the other hand, from (23.2) we obtain $\varrho = W a\, T_*^4$. These relations give $T_1 = W^{\frac{1}{4}}\, T_*$, and this leads to values of T_1 of the order of *ten degrees*. Thus, if the radiation in the nebula remained equilibrium radiation, i. e. for a given integrated density of radiation the energy distribution in the spectrum were given not by formula (23.2) but by Planck's formula with a temperature T_1, the maximum of this distribution would be moved into the far infra-red region of the spectrum.

The above-mentioned constancy of the spectral composition of the radiation which passes from the nuclei to the nebulae makes it possible to ascertain the direction in which the processes of transformation of radiation take place in nebulae. It is known from thermodynamics that, in such cases, the interaction of radiation with matter leads to a redistribution of the energy in frequency, towards the establishment of a more probable distribution. Consequently, the nebulae must so transform the radiation from their nuclei that the long-wave part of the spectrum is strengthened and the short-wave part weakened.

The process of transformation of radiation in nebulae can be characterised by means of Rosseland's Theorem, which we shall now explain. Let us consider an atom which may be in any of three steady states 1, 2, and 3, with corresponding energies $\varepsilon_1 < \varepsilon_2 < \varepsilon_3$. As a result of the absorption of light quanta, there are possible both simple transitions, amounting to resonance scattering, of the kind $1 \to 2 \to 1$, and more complex processes of a cyclic nature. The most important of these are the two converse processes of the kind

$$1 \to 2 \to 3 \to 1\,, \qquad 1 \to 3 \to 2 \to 1\,.$$

The former of these results in the emission of one high-frequency quantum $h\nu_{13}$ as a consequence of the absorption of two quanta of lower frequencies ν_{12} and ν_{23}, and the latter to the partition of one absorbed quantum of frequency ν_{13} into two quanta of frequencies ν_{12} and ν_{23}.

Rosseland's Theorem asserts that, when there is dilution of radiation, i. e. when the density of radiation in various frequencies is given by formula (23.2), cyclic transitions of the kind $1 \to 3 \to 2 \to 1$ take place more often than those of the kind $1 \to 2 \to 3 \to 1$, i. e. quanta of high frequency are transformed into quanta of low frequency more often than the converse.

The number of transitions from the first state to the second in unit volume per unit time is $n_1 B_{12} \varrho_{12}$, where n_1 is the number of atoms in the first state per unit volume, B_{12} is the Einstein absorption coefficient, and ϱ_{12} is the density of radiation of the frequency corresponding to the transition from the first state to the second. Of this number of atoms, a part passes spontaneously or by a stimulated transition back to the first state, and part, having absorbed radiation, passes to the third state. The ratio of the number of transitions from the second state to the third state to the total number of transitions from the second state is

$$B_{23} \varrho_{23}/(A_{21} + B_{21} \varrho_{12} + B_{23} \varrho_{23}) .$$

From the third state, transitions are possible to both the first and second states. We are interested only in transitions from the third state to the first. The proportion of such transitions is

$$(A_{31} + B_{31} \varrho_{13})/(A_{31} + B_{31} \varrho_{13} + A_{32} + B_{32} \varrho_{23}) .$$

Thus we find that in unit time the following number of atoms pass along the path $1 \to 2 \to 3 \to 1$:

$$N_{1\to2\to3\to1} = n_1 B_{12} \varrho_{12} \frac{B_{23} \varrho_{23}}{A_{21} + B_{21} \varrho_{12} + B_{23} \varrho_{23}} \cdot \frac{A_{31} + B_{31} \varrho_{13}}{A_{31} + B_{31} \varrho_{13} + A_{32} + B_{32} \varrho_{23}} . \tag{23.5}$$

We can similarly find the number of atoms which pass per unit time along the path $1 \to 3 \to 2 \to 1$. This number is

$$N_{1\to3\to2\to1} = n_1 B_{13} \varrho_{13} \frac{A_{32} + B_{32} \varrho_{23}}{A_{32} + B_{32} \varrho_{23} + A_{31} + B_{31} \varrho_{13}} \cdot \frac{A_{21} + B_{21} \varrho_{12}}{A_{21} + B_{21} \varrho_{12} + B_{23} \varrho_{23}} . \tag{23.6}$$

For the ratio of the number of processes of the kind $1 \to 2 \to 3 \to 1$ to that of processes of the kind $1 \to 3 \to 2 \to 1$, we obtain from (23.5) and (23.6)

$$\frac{N_{1\to2\to3\to1}}{N_{1\to3\to2\to1}} = \frac{B_{12} \varrho_{12} B_{23} \varrho_{23} (A_{31} + B_{31} \varrho_{13})}{B_{13} \varrho_{13} (A_{32} + B_{32} \varrho_{23}) (A_{21} + B_{21} \varrho_{12})} . \tag{23.7}$$

To simplify this expression, we introduce Einstein's relations

$$A_{ki} = B_{ik} \frac{g_i}{g_k} \sigma_{ik}, \quad B_{ki} = \frac{g_i}{g_k} B_{ik} , \tag{23.8}$$

where

$$\sigma_{ik} = 8 \pi h \nu_{ik}^3/c^3 , \tag{23.9}$$

and g_i, g_k are the corresponding statistical weights. Next, we can write

$$\varrho_{ik} = W \sigma_{ik} \bar{\varrho}_{ik} , \tag{23.10}$$

where

$$\bar{\varrho}_{ik} = \frac{1}{e^{h\nu_{ik}/kT_*} - 1} . \tag{23.11}$$

Using these relations we find instead of (23.7)

$$\frac{N_{1\to 2\to 3\to 1}}{N_{1\to 3\to 2\to 1}} = W \frac{\bar{\varrho}_{12}\,\bar{\varrho}_{23}\,(1 + W\,\bar{\varrho}_{13})}{\bar{\varrho}_{13}\,(1 + W\,\bar{\varrho}_{12})\,(1 + W\,\bar{\varrho}_{23})} . \tag{23.12}$$

When $W = 1$, i. e. in the photosphere of the star, this ratio is exactly equal to unity, as we should expect. As W decreases, so does the ratio. Since W in the nebulae is of the order of 10^{-14}, and the factor $\bar{\varrho}_{12}\,\bar{\varrho}_{23}/\bar{\varrho}_{13}$ has a value of the order of unity, in this case

$$\frac{N_{1\to 2\to 3\to 1}}{N_{1\to 3\to 2\to 1}} \approx W . \tag{23.13}$$

Consequently, in the nebulae we can completely neglect the number of transitions of the kind $1 \to 2 \to 3 \to 1$ compared with that of transitions of the kind $1 \to 3 \to 2 \to 1$. Exactly the same kind of picture can be obtained for atoms with a larger number of levels, in which more complex cyclic processes are possible.

Thus, according to Rosseland's Theorem, the radiation of the nebulae can be explained as the result of the transformation of the ultra-violet quanta from the nuclei into quanta of lower frequency (and, in particular, into quanta in the visible region of the spectrum). Since the nuclei of the nebulae are very hot stars, it is not surprising that the part of the energy, in the ultra-violet region of the spectrum of the nucleus, which is transformed into visible radiation by the nebula exceeds the energy emitted by the nucleus in the visible region of the spectrum. We can thus explain the fact that the nebulae have greater visual luminosities than their nuclei.

3. **The determination of the temperatures of the nuclei from hydrogen lines.** In the last section we have considered a nebula consisting of imaginary atoms with three levels. There we assumed that the nebula is under the influence only of the radiation of the central star, and did not take into account the diffuse radiation of the nebula itself. Let us now consider the radiation of an actual nebula consisting of hydrogen atoms.

First of all, we notice that, owing to the small density of radiation in the nebula, the overwhelming majority of the atoms are in the ground state. If the mass of the nebula is sufficiently great, it will therefore be opaque in the frequencies which are absorbed by the ground state of the hydrogen atom (the Lyman series) and at the same time completely

transparent in the frequencies of the lines of subordinate series (Balmer, Paschen, etc.). Consequently, the nebula will absorb the radiation of the nucleus in the frequencies of the Lyman series and emit, instead of these, the frequencies of the subordinate series (in particular, the observed Balmer series), which pass unimpeded through the nebula.

However, the assumption whereby the energy emitted by the nebula in the Balmer frequencies is obtained from the radiation of the nucleus in the various Lyman lines makes it necessary to assume that the temperatures of the nuclei are very high, over 100,000°. We know, however, that these temperatures, at least for some nuclei, must be about 30,000°. Hence it must be supposed that there are so many hydrogen atoms in the nebula that they absorb not only quanta in the Lyman lines of the star's spectrum, but also quanta in the continuous spectrum beyond the limit of the Lyman series. When these quanta are absorbed, photoionisation of the hydrogen atoms takes place, and later, after recombination, there occur "cascade" transitions of the electrons from one level to another, and quanta are emitted in the lines of the Balmer, Paschen, and other series.

Thus we assume that the optical thickness of the nebula beyond the limit of the Lyman series is greater than unity, or at any rate not small compared with unity. The ratio of the absorption coefficient in the first few lines of the Lyman series to that just beyond the series limit is of the order of 10^4 or 10^5. Hence the optical thickness in these lines must be, on the above assumption, also of the order of 10^4 or 10^5 or more, i. e. it is very large.

Having ascertained the conditions which exist in a nebula, let us investigate more closely what happens to the ultra-violet quanta absorbed by the nebula; by "ultra-violet quanta" or "L_c quanta" we understand quanta with frequencies greater than that of the limit of the Lyman series.

Let an L_c quantum emitted by the star be absorbed by the nebula, causing in it the ionisation of some hydrogen atom. After a time, the free electron will be reunited with some proton. Here there are two possibilities: (1) the electron falls directly into the first level, (2) the electron falls into one of the upper levels. In the first case an L_c quantum is emitted, and the whole process begins again. In the second case the electron performs some chain of transitions, the last of which is a transition to the first level. The dilution of the radiation is so great, and the density of matter in the nebula is so small, that this chain of transitions is not interrupted in the vast majority of cases.

Let us assume that the electron is captured directly into the second level. A quantum in the Balmer continuum is thereby emitted, and this escapes from the nebula. The electron then passes from the second level to the first, emitting a quantum of the first line (α) of the Lyman series.

On the above assumption concerning the very great optical thickness of the nebula in the lines of the Lyman series, this Lyman α quantum, after passing a short distance in the nebula, will be absorbed by some other atom in its normal state. This atom will then be in the second state and, in consequence of the absence of external disturbances, will pass spontaneously to its normal state again, emitting a Lyman α quantum. Thus the Lyman α quantum will undergo only processes of scattering, and this will continue until it reaches the boundary of the nebula and escapes. Consequently, in this case, i.e. for recombination to the second level, there escape from the nebula one quantum in the Balmer continuum and one quantum in the Lyman α line.

Let us now suppose that the electron is captured into the third level. A quantum beyond the limit of the Paschen series is thereby emitted, which escapes from the nebula. The electron then has two possibilities: either to pass directly to the first level, with emission of a Lyman β quantum, or to make a transition first to the second level and then to the first, with successive emission of H_α and Lyman α quanta. However, the Lyman β quantum which arises when the former possibility is realised will necessarily be absorbed in the nebula, in consequence of the latter's great optical thickness in the Lyman lines, and thus an atom in the third state will again be formed. Hence the second of the two possibilities will sooner or later be realised. The H_α quantum thus produced will leave the nebula without hindrance, and the Lyman α quantum, as was explained above, will do so after numerous scatterings. Consequently, when an electron is captured into the third level, three quanta should escape from the nebula: one in the Paschen continuum, one in the H_α line, and one in the Lyman α line.

Similar considerations may be given for recombinations to the higher levels. It is easily seen that, in every case, the chain of transitions of the electron from level to level ends with the formation of a Lyman α quantum, which must be preceded by a transition of the electron to the second level with emission of a Balmer quantum. Thus we reach the following important conclusion: from every L_c quantum absorbed and transformed by the nebula, there are necessarily formed one Lyman α quantum and one Balmer quantum.

If the nebula transforms all the L_c quanta emitted by the star, it follows that the number of these quanta must be equal to the number of Balmer quanta emitted by the nebulae. It is evident that this will hold in the case where the optical thickness of the nebula beyond the limit of the Lyman series is considerably greater than unity. In general, it can be asserted that the number of Balmer quanta emitted by the nebula does not exceed the number of L_c quanta emitted by the central star, i.e.

$$N_B \leq N_{L_c}{}^* . \tag{23.14}$$

On this theorem is based the method, proposed by H. ZANSTRA [181], of determining the colour temperatures of the nuclei of planetary nebulae. Since the number of Balmer quanta emitted by the nebula gives a lower limit to the number of L_c quanta emitted by the nucleus, if we compare the number of quanta emitted by the nebula in the Balmer lines with the number of quanta emitted by the star in the visible part of the spectrum, we are in fact comparing some lower limit to the intensity of the nucleus spectrum beyond the Lyman series limit with the intensity of the nucleus spectrum in the visible region. This makes it possible to define a lower limit to the temperature.

If the nucleus emits according to Planck's Law, and its radius is r_* and its temperature T_*, the total number of L_c quanta emitted by it in unit time is

$$N_{L_c}{}^* = 4\,\pi\,r_*{}^2\,\frac{2\,\pi}{c^2}\int\limits_{\nu_0}^{\infty}\frac{\nu^2\,d\nu}{e^{h\nu/kT_*}-1}\,, \tag{23.15}$$

where ν_0 is the frequency of the Lyman series limit.

On the other hand, let E_i denote the total amount of energy emitted per second by the whole nebula in the ith Balmer line, and let $(\partial E_*/\partial \nu)_i$ be the total amount of energy emitted by the nucleus in unit time and unit frequency interval at the same point in the spectrum. The dimensionless ratios

$$A_i = \frac{E_i}{\nu_i(\partial E_*/\partial \nu)_i} \tag{23.16}$$

can be determined directly from observation. But

$$\left(\frac{\partial E_*}{\partial \nu}\right)_i = 4\,\pi\,r_*{}^2\,\frac{2\,\pi\,h\,\nu_i^3}{c^2}\,\frac{1}{e^{h\nu_i/kT_*}-1}\,. \tag{23.17}$$

Hence the total number of Balmer quanta emitted by the nebula is

$$N_B = \sum\frac{E_i}{h\,\nu_i} = \frac{1}{h}\sum A_i\left(\frac{\partial E_*}{\partial \nu}\right)_i = 4\,\pi\,r_*{}^2\,\frac{2\,\pi}{c^2}\sum\frac{\nu_i^3\,A_i}{e^{h\nu_i/kT_*}-1}\,. \tag{23.18}$$

From the inequality (23.14) we obtain, by (23.15) and (23.18),

$$\sum\frac{\nu_i^3\,A_i}{e^{h\nu_i/kT_*}-1} \le \int\limits_{\nu_0}^{\infty}\frac{\nu^2\,d\nu}{e^{h\nu/kT_*}-1}\,. \tag{23.19}$$

We write

$$h\,\nu/k\,T_* = x\,,\quad h\,\nu_0/k\,T_* = x_0\,,\quad h\,\nu_i/k\,T_* = x_i\,. \tag{23.20}$$

Then we have instead of (23.19)

$$\sum \frac{x_i^3 A_i}{e^{x_i} - 1} \leq \int_{x_0}^{\infty} \frac{x^2 \, dx}{e^x - 1} . \tag{23.21}$$

The summation on the left-hand side of this inequality is extended not only over all the lines of the Balmer series, but also over the Balmer continuum.

The inequality (23.21) is solved by trial. Taking various T_* and calculating x_0 and x_i, we find that for some value of T_* we have an equality. A lower limit to the temperature of the nucleus is thus determined. Tables which facilitate the solution of this problem are given in the work of ZANSTRA, and graphs in the work of B. A. VORONTSOV-VEL'YAMINOV [173].

The method given for determining the temperatures of the nuclei of planetary nebulae has been applied in practice by ZANSTRA and by some other authors. As an example, we give the results of ZANSTRA, obtained from his own measurements of the quantities A_i.

Nebula	NGC 6543	$T_* = 39{,}000°$
	NGC 6572	$T_* = 40{,}000°$
	NGC 7009	$T_* = 55{,}000°$.

It must be noticed that the method is very insensitive to small errors in the determination of the A_i, since we are essentially comparing two parts of the nucleus spectrum which are very distant from each other.

The radiation of the planetary nebulae in the lines of other atoms (but not, as we shall see below, in all lines) occurs in the same way as in the hydrogen lines, as a result of photo-ionisation under the action of the ultra-violet radiation of the star and subsequent recombinations. In particular, the nebulae radiate thus in the lines of helium and of ionised helium. From the intensity of these lines we can determine the temperatures of the nuclei, just as from the intensity of the hydrogen lines. Such determinations again lead to very high values of the temperatures of the nuclei. Here it has been found that the temperatures of the same star, as found from lines of different atoms, differ fairly considerably. For example, for the nucleus of the nebula NGC 7009 ZANSTRA obtained $T_* = 70{,}000°$ from ionised helium and $T_* = 55{,}000°$ from hydrogen. In many cases the discrepancy is even more marked.

To explain this discrepancy, several reasons may be adduced. We shall give only two of them, the most important ones.

(1) ZANSTRA's method does not give the actual temperature of the star, but only a lower limit to it. If the nebula absorbs only a small part of the energy from the star beyond the principal series limit of the atom concerned, this lower limit to the temperature may lie considerably below the actual temperature. This is apparently so, in some cases, for the temperatures found from the hydrogen lines, since, in consequence of the strong ionisation of hydrogen in the nebulae, the absorption of the L_c radiation of the nucleus may be incomplete.

(2) It may be that the energy distribution in the spectrum of the nucleus does not obey Planck's Law, but deviates considerably from it. In this case the concept of the colour temperature has no unambiguous significance.

4. The determination of the temperatures of the nuclei from "nebulium" lines. We have said above that the majority of lines in the spectra of the planetary nebulae arise as a result of photo-ionisation and subsequent recombination. However, it is easy to show that some lines, including the principal nebular lines N_1 and N_2, cannot arise in this way.

Let us assume, in fact, that the radiation of the nebula in the N_1 and N_2 lines occurs at the expense of the energy from the star beyond the limit of the principal series of doubly ionised oxygen. In this case the number of quanta emitted by the star beyond the limit of this series should be not less than the number of quanta emitted by the nebula in the N_1 and N_2 lines. But the limit of the principal series of the O III atom is in the very far ultra-violet region of the spectrum (the ionisation potential of O III is approximately four times the ionisation potential of hydrogen). It follows from this that, if our assumption were correct, the temperature of the star would be extremely high — in some cases, upwards of a million degrees.

To this we may add the following consideration. Since the ionisation potentials of O III and He II are almost the same (54·5 and 54·2 eV respectively), these atoms absorb the energy from the star in the same region of the spectrum. Hence, if the nebula radiated in the N_1 and N_2 lines only at the expense of this energy, the N_1 and N_2 lines would not exceed in intensity the lines of He II (since there are at least as many helium atoms as oxygen atoms). In reality, there are a number of nebulae in whose spectra the lines of He II are almost unnoticeable, whereas the N_1 and N_2 lines are very intense.

These facts show that there must exist in the nebulae some mechanism of excitation of the atoms, other than recombinations, and I. S. BOWEN has indicated such a mechanism. He pointed out that the excitation potentials of the levels from which the "nebulium" lines are emitted are very small (for example, only 2·5 eV for the N_1 and N_2 lines). Consequently, a considerable fraction of the free electrons in the nebula

must have energies sufficient to excite these levels by collision. For this reason, BOWEN supposed that the nebulae radiate in the "nebulium" lines at the expense of *the kinetic energy of free electrons.*

It is clear that the energy of the free electrons is obtained ultimately from the ultra-violet radiation of the star. This radiation not only detaches the electrons from the atoms, but endows them with some kinetic energy. A part of this energy of the electron is expended on exciting the "nebulium" lines, before it is captured by an ion.

The ideas just given on the radiation of nebulae in the "nebulium" lines enabled ZANSTRA to give another method of determining the temperatures of the nuclei.

We shall suppose that the free electrons are formed mainly by the ionisation of hydrogen atoms. If the ionisation takes place by the absorption of a quantum of frequency ν, the detached electron will acquire a kinetic energy of

$$\tfrac{1}{2}\, m\, v^2 = h\,\nu - h\,\nu_0\,,$$

where ν_0 is the ionisation frequency of hydrogen. The total number of quanta emitted by the star in the frequency interval from ν to $\nu + d\nu$ in 1 second is

$$4\,\pi\, r_*^2\, \frac{2\,\pi\,\nu^2}{c^2}\, \frac{d\nu}{e^{h\nu/kT_*} - 1}\,.$$

Hence the total amount of kinetic energy acquired by free electrons in 1 second, if all the L_c quanta emitted by the star are absorbed, is

$$4\,\pi\, r_*^2\, \frac{2\,\pi\, h}{c^2} \int_{\nu_0}^{\infty} \frac{(\nu - \nu_0)\,\nu^2}{e^{h\nu/kT_*} - 1}\, d\nu\,. \tag{23.22}$$

On the other hand, the energy emitted by the nebula in the "nebulium" lines can be represented in the form

$$4\,\pi\, r_*^2\, \frac{2\,\pi\, h}{c^2} \sum_{\text{nebulium}} \frac{\nu_i^4}{e^{h\nu_i/kT_*} - 1}\, A_i\,, \tag{23.23}$$

where the A_i are the quantities determined from observation by formula (23.16), and the summation is taken over all the "nebulium" lines which are excited as a result of collisions with electrons.

Since the energy expended on the excitation of the "nebulium" lines cannot exceed the kinetic energy acquired by the electrons, we find by comparing (23.22) and (23.23)

$$\sum_{\text{nebulium}} \frac{\nu_i^4\, A_i}{e^{h\nu_i/kT_*} - 1} \leq \int_{\nu_0}^{\infty} \frac{(\nu - \nu_0)\,\nu^2}{e^{h\nu/kT_*} - 1}\, d\nu\,, \tag{23.24}$$

or, using the notation (23.20),

$$\sum_{\text{nebulium}} \frac{x_i^4}{e^{x_i} - 1} A_i \leq \int_{x_0}^{\infty} \frac{(x - x_0)\, x^2}{e^x - 1} \, \mathrm{d}x \,. \tag{23.25}$$

In the inequality (23.25), the unknown is the temperature T_* which appears in both x_0 and x_i. On solving this inequality, we find some lower limit to the temperature of the star.

The results of this determination of lower limits of temperature for the nuclei of three nebulae are

NGC 6543	$T_* = 37{,}000°$
NGC 6572	$T_* = 38{,}000°$
NGC 7009	$T_* = 50{,}000°$.

We see that the results obtained by this method agree fairly well with those obtained from hydrogen lines. It may be that this close agreement of the results obtained by the two methods is due to the fact that they not only are lower limits, but are actually close to the temperatures sought.

It should be mentioned that the finding of the quantities A_i from observation presents considerable difficulties. A modification of the above method is therefore of great practical importance; it is based on the fact that the N_1 and N_2 lines account for most of the visual luminosity of the nebula. Moreover, as we have seen, these lines characterise to some extent the intensity of the radiation from the star beyond the limit of the Lyman series. Hence we can estimate the temperature of the star from the ratio of the visual luminosity of the nebula to the visual (or photographic) luminosity of the star, i.e. from the difference $m_* - m_n$. It is evident that, the greater this difference, the higher the temperature.

In this way ZANSTRA has determined the temperatures of the central stars of a fairly large number of nebulae. He has found that, in some cases, the temperature of the central star reaches 100,000° or more. The high temperatures obtained by this method are usually confirmed by other indications, and in particular by the intensities of the lines of He II.

Chapter 24. The physical state of matter in the nebulae

1. The conditions necessary for the appearance of forbidden lines. Let us now turn to a more detailed discussion of the fundamental fact that very intense forbidden lines are observed in the spectra of planetary nebulae.

Besides the principal nebular lines N_1 and N_2, which belong to the O III atom, forbidden lines of the atoms O I, O II, N II, S II and others are also visible in the spectra of nebulae. A list of the brightest of these lines is given in Table 17.

Table 17

Wavelength	Transition	Atom	Excitation potential
3726	${}^4S^0_{3/2} - {}^2D^0_{3/2}$	O II	3·31
3729	${}^4S^0_{3/2} - {}^2D^0_{5/2}$	O II	3·31
4363	${}^1D_2 - {}^1S_0$	O III	5·33
4959	${}^3P_1 - {}^1D_2$	O III	2·50
5007	${}^3P_2 - {}^1D_2$	O III	2·50
6583	${}^3P_2 - {}^1D_2$	N II	1·89

As an example, we give in Fig. 60 the energy level diagram of the O III atom. Together with other transitions, the figure shows the forbidden transitions corresponding to the lines N_1, N_2 and 4363 Å.

The presence of forbidden lines in the spectra of nebulae is not, in itself, surprising. The transition probabilities for forbidden lines are obtained as zero only by an approximate calculation. In fact they are different from zero, although comparatively small. Whereas the Einstein spontaneous transition probabilities are of the order of 10^8 per second for permitted lines in the visible part of the spectrum, for forbidden lines they are less by factors of a million or a thousand million. For example, for the N_1 and N_2 lines the spontaneous transition probabilities are 0·018 and 0·006 per second respectively.

However, the forbidden lines are not observed in ordinary stellar spectra. We must therefore ascertain what is the difference between the conditions in the nebulae and those in stellar atmospheres which leads to the appearance of forbidden lines in the spectra of the nebulae.

First of all, we must notice that those forbidden lines which start from upper levels from which permitted transitions to lower levels are also possible cannot reach a high relative intensity, since from the upper state concerned permitted transitions will occur millions of times more often than forbidden ones. Hence the forbidden lines can become comparatively intense only when *the upper state is metastable*, i.e. no transitions from it to lower energy levels are possible, apart from the forbidden ones.

However, whereas the atom remains in ordinary excited states for a time of the order of 10^{-8} second, *an atom can remain for several seconds or more in a metastable state.* Thus, for example, the mean lifetime of an O III atom in the $^1D_2$ state, from which the N_1 and N_2 lines are emitted, is 42 seconds. Consequently, in order that the forbidden lines should be able to be emitted from the metastable states, it is necessary

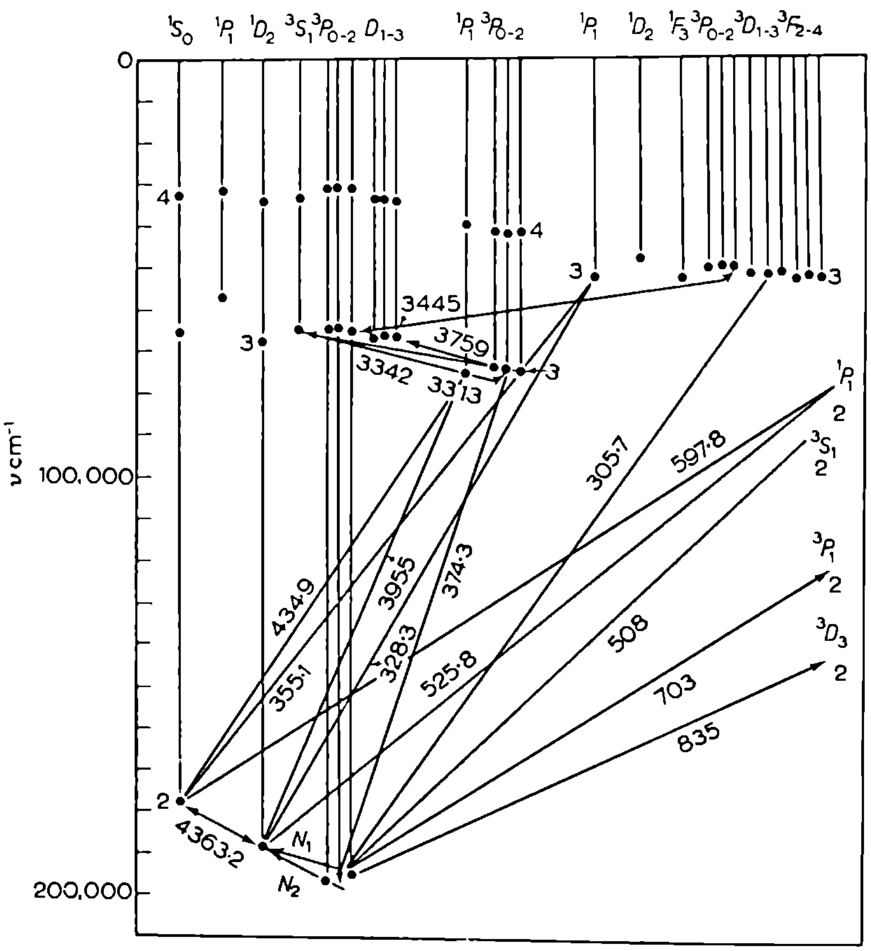

Fig. 60

that the atom should not undergo external perturbations during a very long interval of time. In particular, the atom must not experience frequent encounters with free electrons, since these can transfer it from the metastable state upwards by a collision of the first kind or downwards, without the emission of the forbidden line, by a collision of the second kind. Similarly, the atom must not be subjected to strong action of radiation, since it can pass upwards from the metastable state by absorbing a light quantum. Thus the long lifetime of an atom in the metastable state requires some restrictions to be placed on the density of matter and of radiation; if these are not fulfilled, the forbidden line

cannot appear. In other words, for the appearance of forbidden lines in the spectrum of any object, it is necessary that the density of matter and of radiation in it should be sufficiently small.

The absence of forbidden lines in stellar spectra indicates that this condition is not fulfilled in the atmospheres of stars. This means that an atom which has entered the metastable state is rapidly removed from it by the action of light quanta or of free electrons, which exist in large numbers in the atmospheres of stars, and so the forbidden line is not emitted. Conversely, the presence of numerous very intense forbidden lines in the spectra of the planetary nebulae points to an extremely low density of radiation and of matter in these objects. In the following section the conditions necessary for the appearance of forbidden lines will be considered quantitatively.

2. **The accumulation of atoms in metastable states.** Since the conditions in nebulae are such that atoms which have entered a metastable state are able to remain in it for a fairly long time (until they make a spontaneous downward transition), a very large number of atoms must accumulate in metastable states. For this reason alone are bright forbidden lines emitted, since the intensity of a line is proportional to the number of atoms in the initial state, and to the probability of the corresponding spontaneous transition, and the probabilities of spontaneous transitions from metastable states are extremely small.

The theoretical problem of the accumulation of atoms in metastable states was first solved by V. A. AMBARTSUMYAN [3], and we shall give here the results obtained by him.

Let us consider, for simplicity, an atom having three energy levels. We assume at first that the excitation of the atom is caused by radiation only. In the steady state, the number of atoms in each of the levels should be constant. Hence we have

$$\left.\begin{aligned} n_1 B_{12}\varrho_{12} + n_1 B_{13}\varrho_{13} &= n_2 A_{21} + n_3 A_{31}\,; \\ n_1 B_{13}\varrho_{13} + n_2 B_{23}\varrho_{23} &= n_3 A_{31} + n_3 A_{32}\,. \end{aligned}\right\} \tag{24.1}$$

Here we have neglected stimulated transitions, since we are assuming that the radiation is strongly diluted ($W \ll 1$). From the equations (24.1), we obtain by eliminating n_3

$$\frac{n_2}{n_1} = \frac{B_{12}\varrho_{12} + (1-p)\,B_{13}\varrho_{13}}{A_{21} + pB_{23}\varrho_{23}}\,, \tag{24.2}$$

where $p - A_{31}/(A_{31}+A_{32})$. We use the relations between the Einstein transition coefficients, and the notations (23.10) for the quantities ϱ_{ik}. We find for the ratio n_2/n_1, instead of (24.2),

$$\frac{n_2}{n_1} = W\frac{A_{21}\,(g_2/g_1)\,\bar{\varrho}_{12} + (1-p)\,A_{31}\,(g_3/g_1)\,\bar{\varrho}_{13}}{A_{21} + pA_{32}\,(g_3/g_2)\,W\bar{\varrho}_{23}}\,. \tag{24.3}$$

We shall now suppose that the second state is metastable, i.e. $A_{21} \ll A_{31}$ and A_{32}. Then the first term in the numerator of formula (24.3) can be neglected in comparison with the second. However, either of the two terms in the denominator may be the greater, depending on the circumstances. Hence we shall analyse two possible cases.

I. $W \ll A_{21}/A_{32}$. In this case we can omit the second term in the denominator, and we have

$$\frac{n_2}{n_1} = W(1-p)\frac{A_{31}}{A_{21}}\frac{g_3}{g_1}\bar{\varrho}_{13}. \tag{24.4}$$

II. $W \gg A_{21}/A_{32}$. In this case, on the other hand, we can omit the first term in the denominator, and we therefore have

$$\frac{n_2}{n_1} = \frac{g_2}{g_1}\frac{\bar{\varrho}_{13}}{\bar{\varrho}_{23}} \approx \frac{g_2}{g_1}\bar{\varrho}_{12}. \tag{24.5}$$

Let us also consider, for comparison, the case where the second level is not metastable, i.e. the transition $2 \to 1$ is permitted. Denoting by $A_{21}{}^0$ the coefficient A_{21} for the permitted transition, we can write $A_{21}{}^0 \approx A_{31}$, A_{32}, and we find from formula (24.3)

$$\frac{n_2{}^0}{n_1} = W\frac{g_2}{g_1}\bar{\varrho}_{12} + W(1-p)\frac{A_{31}}{A_{21}{}^0}\frac{g_3}{g_1}\bar{\varrho}_{13}. \tag{24.6}$$

It is evident that the two terms on the right-hand side of this formula can be considered to be of the same order of magnitude.

On comparing formulae (24.4) and (24.6), we see that, in the metastable state in the first of the cases considered above, the ratio n_2/n_1 exceeds its value for an ordinary excited state by the same factor as the probability of the permitted transition exceeds that of the forbidden transition ($n_2/n_2{}^0 \approx A_{21}{}^0/A_{21}$). Consequently, in this case the intensities of the forbidden lines attain the same order of magnitude as those of the permitted lines (since $n_2 A_{21} \approx n_2{}^0 A_{21}{}^0$).

In the second case considered, however, the ratio n_2/n_1 is approximately determined by Boltzmann's formula, and, as is shown by a comparison of formulae (24.5) and (24.6), it is in order of magnitude $1/W$ times the value of this ratio for an ordinary excited level. Consequently, a very marked accumulation of atoms in the metastable state occurs in this case also. However, the number of transitions from the metastable state does not equal in order of magnitude the number of transitions from the ordinary state. In fact, we have

$$n_2 A_{21} \approx n_1 \frac{g_2}{g_1}\bar{\varrho}_{12} A_{21} \approx n_2{}^0 \frac{A_{21}}{W} \ll n_2{}^0 A_{21}{}^0.$$

This means that, in case II, the intensity of the forbidden line is small in comparison with that of the permitted line.

The physical significance of cases I and II is as follows. The first term in the denominator of formula (24.3) corresponds to forbidden transitions $2 \to 1$, and the second to transitions upwards from the metastable state ($2 \to 3$). In the first case the transitions $2 \to 1$ predominate, and the forbidden line is seen at full strength. In the second case, in consequence of the comparatively high density of radiation, the transitions $2 \to 3$ predominate, and the forbidden line is faint.

Since in the planetary nebulae W is of the order of 10^{-13}, almost all metastable states satisfy the conditions of the first case. It may be that the only exception is the 2^3S state of helium, which has an extremely long lifetime. On the other hand, in envelopes of small radius (i.e. those of Wolf-Rayet, Be, etc., type stars), almost all the metastable states belong to case II.

The above shows the state of affairs concerning the accumulation of atoms in metastable states when the excitation is by the action of radiation. However, as was found above, the brightest of the forbidden lines in the spectra of nebulae are excited not by radiation, but by collisions with free electrons. We shall therefore consider also the problem of the population of the metastable states for this second mechanism of excitation.

Let us take an atom with two energy levels. Let $n_1 b_{12}$ be the number of transitions from the normal to the metastable state in collisions of the first kind, $n_2 a_{21}$ the number of transitions from the metastable to the normal state in collisions of the second kind, and $n_2 A_{21}$ the number of spontaneous transitions from the metastable to the normal state (all per unit volume and unit time). The condition for a steady state gives

$$n_1 b_{12} = n_2(a_{21} + A_{21}) \,. \tag{24.7}$$

The coefficients b_{12} and a_{21} are proportional to the density of free electrons, and depend on their velocity distribution. We shall suppose that the velocities of the free electrons are distributed according to Maxwell's Law*. In this case the following relation exists between the coefficients b_{12} and a_{21}:

$$b_{12} = a_{21} \frac{g_2}{g_1} e^{-h\nu_{12}/kT_e} \,, \tag{24.8}$$

* It can be shown that, in the time between ionisation and recombination, each electron undergoes a very large number of collisions with other electrons. Hence this assumption is not suspect.

where T_e is the temperature of the electron gas*. Hence we obtain from equation (24.7)

$$\frac{n_2}{n_1} = \frac{a_{21}}{A_{21}+a_{21}} \frac{g_2}{g_1} e^{-h\nu_{12}/kT_e} . \tag{24.9}$$

If $a_{21} \gg A_{21}$, formula (24.9) becomes Boltzmann's formula. However, this inequality means that stimulated transitions predominate over spontaneous ones. In this case, therefore, the forbidden line will be faint or not visible at all. If the converse inequality holds, the majority of atoms which are transferred from the normal to the metastable state in a collision of the first kind will return spontaneously, emitting a quantum in the forbidden line. Hence, if these collisions are sufficiently numerous, the forbidden line will be very strong.

We shall see below that the density of matter in planetary nebulae is extremely small ($n_e \approx 10^4$). The inequality $a_{21} \ll A_{21}$ therefore holds for these nebulae, i.e. collisions of the second kind occur rarely and cannot "quench" the forbidden line. Collisions of the first kind, on the other hand, occur sufficiently frequently to transfer a fairly large number of atoms from the normal to the metastable state. This difference between the numbers of the two kinds of collisions is due to the fact that an atom spends much less time in an excited state (even if metastable) than in the normal state.

From what has been said in the present section, we can draw the following conclusion. In order that the forbidden lines should be comparable in intensity with the permitted lines, it is necessary that two conditions should be fulfilled:

$$(1) \quad W \ll A_{21}/A_{21}^0 ,$$

$$(2) \quad a_{21} \ll A_{21} .$$

These determine the upper limits that may be reached by the density of radiation and of matter. It is evident that both these conditions must be fulfilled, whatever is the mechanism of excitation of the atoms.

3. The determination of the temperatures of the nebulae. The fact that the metastable states of the O III ion are excited by electron collisions enabled V. A. AMBARTSUMYAN to propose a very simple method of determining the electron temperatures of the nebulae. He pointed out that the O III ion has not only the metastable level $^1D_2$, from which the N_1 and N_2 lines start, but also the higher metastable level $^1S_0$; the line 4363 Å is emitted in the transition from this level to the

* To derive this relation, we may consider the state of thermodynamic equilibrium. In this case we must have $n_1 b_{12} = n_2 a_{21}$ and $n_2/n_1 = (g_2/g_1)\, e^{-h\nu_{12}/kT}$. Formula (24.8) follows from these two relations; it is valid, of course, in all cases where there is a Maxwellian velocity distribution of electrons, and not only in thermodynamic equilibrium.

$^1D_2$ level (see Fig. 60). The excitation potentials of these levels are 2·5 and 5·3 eV respectively. It is evident that, the higher the electron temperature of the nebula, the greater will be the ratio of the number of atoms in the $^1S_0$ state to that in the $^1D_2$ state, and consequently, the brighter will be the line 4363 Å in comparison with the N_1 and N_2 lines. Thus we can estimate the electron temperature of the nebula from observations of the relative intensities of the lines 4363 Å and $N_1 + N_2$.

To derive the required formula, we denote the number of O III atoms in the normal and two metastable states by n_1, n_2 and n_3 respectively. Since, in the case of the nebulae, we can neglect transitions from the excited states under the action of collisions, in comparison with the spontaneous transitions, we obtain as the conditions for a steady state

$$\left.\begin{aligned} n_1 b_{12} + n_3 A_{32} &= n_2 A_{21}\,, \\ n_1 b_{13} &= n_3 (A_{31} + A_{32})\,. \end{aligned}\right\} \tag{24.10}$$

It is found that for the O III ion the transition $3 \to 1$ is "forbidden" much more strongly than the transition $3 \to 2$, i.e. $A_{31} \ll A_{32}$. Hence we have from equations (24.10)

$$\frac{n_2 A_{21}}{n_3 A_{32}} = 1 + \frac{b_{12}}{b_{13}}\,.$$

This gives for the required intensity ratio of the lines $N_1 + N_2$ and 4363 Å

$$\frac{E_{21}}{E_{32}} = \frac{\nu_{12}}{\nu_{23}} \left(1 + \frac{b_{12}}{b_{13}}\right). \tag{24.11}$$

Using now the relation (24.8) between the probabilities of collisions of the first and second kinds, we obtain instead of (24.11)

$$\frac{E_{21}}{E_{32}} = \frac{\nu_{12}}{\nu_{23}} \left(1 + \frac{g_2}{g_3}\,\frac{a_{21}}{a_{31}}\, e^{h\nu_{23}/kT_e}\right). \tag{24.12}$$

In passing from formula (24.11) to formula (24.12) we have explicitly introduced the dependence of E_{21}/E_{32} on the electron temperature, since the coefficients a_{21} and a_{31} are almost independent of T_e (because collisions of the second kind, unlike those of the first kind, can be effected by an electron with any velocity). We can suppose that $a_{21}/a_{31} \approx 1$. We have also $g_2/g_3 = 5$. Hence we find approximately, instead of (24.12),

$$E_{N_1+N_2}/E_{4363} = 4{\cdot}5\, e^{33{,}000/T_e}\,. \tag{24.13}$$

This formula serves to determine T_e from the ratio $E_{N_1+N_2}/E_{4363}$ found from observation.

In the spectra of planetary nebulae, the ratio of intensities of the lines $N_1 + N_2$ and 4363 Å varies within fairly wide limits. However, since T_e enters formula (24.13) in an exponent, the values of T_e found from this formula for different nebulae differ little among themselves. According to the determination by D. H. Menzel and his co-workers [84], the electron temperatures of the great majority of the planetary nebulae lie between 6000° and 10,000°.

Another method of determining the electron temperatures of the nebulae is obtained from a consideration of the energy balance of the free electrons [152]. In photo-ionisation, the electrons acquire some kinetic energy. We know that they expend a considerable part of this energy on the excitation of atoms by collision. The other part is transformed into radiation in the continuous spectrum by recombinations and hyperbolic transitions. Since the nebulae are in a steady state, the kinetic energy of the electron gas should remain constant, i.e. the energy acquired by the electrons should equal the energy they lose. This condition determines the temperature of the electron gas. As an upper limit (taking into account collisions of the electrons with O III and hydrogen atoms), temperatures of the order of 9000° to 14,000° are obtained for a number of nebulae.

Thus the electron temperatures of the nebulae are much lower than the temperatures of the central stars. The cooling of the electron gas is chiefly due to collisions with atoms which have levels with low excitation potentials (particularly O III ions). It is these atoms which control the electron temperatures of the nebulae, operating as an unusual "thermostat".

4. The intensities of the Balmer lines. Formula (24.6), which determines the number of atoms in excited (not metastable) states in the nebulae, is only approximate, since we have assumed, in deriving it, that an atom has only three energy levels. However, the conditions in the nebulae are so simple that no difficulty is encountered in calculating the populations of the excited levels for actual atoms. The chief reason for the easiness of such a calculation is the complete transparency of the nebulae to radiation in the lines of subordinate series. Because of this, the population of the excited levels is determined entirely by recombinations (or by collisions in the case of low energy levels), and by the subsequent cascade transitions of the electrons from level to level.

The transparency of the nebulae to radiation in the lines of subordinate series also makes it easy to determine the amount of energy emitted by the nebula in these lines. A comparison of the calculated line intensities with observation forms a good test of the correctness of the theory.

Let us calculate the population of the excited levels of a hydrogen atom. This enables us, in particular, to determine the relative intensities of the Balmer lines (the so-called *Balmer decrement*).

We know that the hydrogen spectrum of the nebulae is due to recombinations. Hence we must first of all obtain an expression for the number of recombinations. It is evident that this number is proportional to the concentrations of ions and of free electrons, i.e. to the quantities n_+ and n_e, and depends on the temperature T_e of the free electrons. Consequently, the number of captures into the ith level is equal to $n_e\, n_+\, C_{if}(T_e)$. The coefficient $C_{if}(T_e)$ is calculated from (20.10). In doing so, we neglect the term $c^3\, \varrho_\nu / 8\, \pi\, h\, \nu^3$ (i.e. the stimulated emission); we take the value of $(\beta_\nu)_k$ from (5.55) and the expression for k'_ν from (5.24). For hydrogen, $Z = 1$, and the factor g' can be taken as unity. The final result is

$$C_{if}(T_e) = \frac{2^9\, \pi^5}{(6\, \pi)^{3/2}}\, \frac{e^{10}}{m^2\, c^3\, h^3} \left(\frac{m}{k T_e}\right)^{3/2} \frac{1}{i^3}\, e^{\chi_i / k T_e}\, E_1 \left(\frac{\chi_i}{k\, T_e}\right), \qquad (24.14)$$

where χ_i is the ionisation potential from the ith state, and the other symbols have their usual meanings.

The calculation of the population of the excited levels is based on the conditions for a steady state, which are that the number of atoms entering any given state must be exactly equal to the number of atoms leaving it.

Let us write down the steady state conditions for the ith level of hydrogen. The number of atoms entering the ith state is the sum of three terms: (1) the number of captures directly into the ith level, $n_e\, n_+\, C_{if}(T_e)$, (2) the number of spontaneous transitions from higher discrete states, $\sum\limits_{k=i+1}^{\infty} n_k\, A_{ki}$, and (3) the number of transitions from the first to the ith state by the absorption of quanta in the Lyman lines, $n_1\, B_{1i}\, \varrho_{1i}$ (we recall that the nebulae are opaque to radiation in the lines of a principal series). Consequently, the total number of transitions to the ith state per unit time and volume is

$$n_e\, n_+\, C_{if}(T_e) + \sum_{k=i+1}^{\infty} n_k\, A_{ki} + n_1\, B_{1i}\, \varrho_{1i}\,.$$

On the other hand, the number of atoms leaving the ith state is

$$n_i \sum_{k=1}^{i-1} A_{ik}\,,$$

since only spontaneous transitions downwards are possible from the ith state. Equating the last two expressions, we obtain

$$n_i \sum_{k=1}^{i-1} A_{ik} = n_e\, n_+\, C_{if}(T_e) + \sum_{k=i+1}^{\infty} n_k\, A_{ki} + n_1\, B_{1i}\, \varrho_{1i} \qquad (i=2,3,4,\ldots)\,. \tag{24.15}$$

We have seen, however, that all quanta in lines of the Lyman series that are emitted by the nebula are absorbed in it again. Consequently, the number of transitions $i \to 1$ is almost exactly equal to the number of transitions $1 \to i$, i.e.

$$n_i\, A_{i1} = n_1\, B_{1i}\, \varrho_{1i}\,. \tag{24.16}$$

Hence we find, instead of (24.15),

$$n_i \sum_{k=2}^{i-1} A_{ik} = n_e\, n_+\, C_{if}(T_e) + \sum_{k=i+1}^{\infty} n_k\, A_{ki} \quad (i=3,4,\ldots)\,. \tag{24.17}$$

Thus we have arrived at a system of linear algebraic equations for the numbers $z_k = n_k/n_e\, n_+$. The solution of this system determines the population of the energy levels of hydrogen (starting from the third).

If the numbers z_k are known, it is easy to find the relative intensities of the emission lines. The energy emitted by the nebula in the line which corresponds to the transition $k \to i$ is

$$E_{ki} = A_{ki}\, h\, \nu_{ik} \int n_k\, \mathrm{d}V\,, \tag{24.18}$$

where the integration is extended over the whole volume of the nebula. But $n_k = z_k\, n_e\, n_+$ and, if we suppose that the electron temperature is constant in the nebula, the numbers z_k can be taken outside the integral sign. We thus obtain

$$E_{ki} = z_k\, A_{ki}\, h\, \nu_{ik} \int n_e\, n_+\, \mathrm{d}V\,. \tag{24.19}$$

For a given T_e, the formula thus found gives the line intensities apart from a constant factor. In particular, for $i = 2$ it determines the Balmer decrement.

The system of equations (24.17) has been approximately solved by G. G. Cillié, who used the first twelve equations ($i=3,4,\ldots,14$) and rejected the remainder. The numbers $10^{20}\, z_k$ which he obtained for various values of T_e are given in Table 18.

Table 18

k \ T_e	5000°	10,000°	20,000°
3	0·41	0·23	0·13
4	0·59	0·33	0·17
5	0·89	0·48	0·24
6	1·29	0·68	0·34
7	1·78	0·92	0·45

More accurate solutions of the system (24.17) have been obtained by D. H. MENZEL and J. G. BAKER [15]. In their tables the quantities b_i are given, which are defined by a relation similar to (20.14):

$$n_i = b_i \, n_e \, n_+ \frac{h^3 \, i^2}{(2\,\pi\, m\, kT_e)^{3/2}} \, e^{\chi_i/kT_e} \,,$$

i.e. they show by what factor the value of $n_i/n_e\, n_+$ in the nebulae differs from its value in a state of thermodynamic equilibrium at temperature T_e. It has been found that the values of the b_i are fairly close to unity, and $b_i \to 1$ as $i \to \infty$, as we should expect. For the lower levels, the results of MENZEL and BAKER do not differ greatly from those of CILLIÉ.

The Balmer decrement, calculated by means of the numbers z_k taken from Table 18, is given in Table 19. From this it is seen that the Balmer decrement depends very little on the electron temperature, and may be regarded as practically constant.

Table 19

T_e	5000°	10,000°	20,000°	observed
H_α	2·70	2·78	2·88	2·77
H_β	1·00	1·00	1·00	1·00
H_γ	0·51	0·50	0·48	0·50
H_δ	0·30	0·29	0·27	0·26
H_ε	0·19	0·18	0·17	0·18

A comparison of theory and observation at first showed some discrepancy between them. The observed Balmer decrement was steeper than the calculated values, and also varied markedly from one nebula to another. However, this discrepancy was later explained by the selective absorption of light in the Galaxy. This brings about a reddening of distant objects, owing to which the observed intensity ratio of the lines H_α/H_β appears greater than it really is. G. A. SHAĬN [136] first pointed out the possibility of such an explanation; he had found a corre-

lation between the Balmer decrement and the galactic latitude of the nebula. Later a correlation was found between the Balmer decrement and the distance of the nebula.

After the absorption of light in the Galaxy had been taken into account, the Balmer decrement was found to be in good agreement with the calculated values. This is seen, for example, from Table 19, in the last column of which is given the observed Balmer decrement, taking account of the absorption of light, averaged over 17 nebulae.

It must be remarked also that the ratio, obtained from observation, of the number of quanta in the Balmer continuum to that in the H_β line deviates from the obvious theoretical value

$$n_e\, n_+\, C_{2f}(T_e)/n_4\, A_{42}$$

in the same sense as the ratio of the line intensities, i.e. the observed ratio is less than the theoretical. Here, of course, the selective absorption of light in space again plays some part. However, the observational data are as yet not very reliable. In particular, the problem is complicated by the presence, at least in some planetary nebulae, of a continuous spectrum, of unknown origin, throughout the visible region*.

5. **Ionisation in the nebulae.** Having found the degree of excitation in the nebulae, let us now determine the degree of ionisation. To do so, we must use the condition of equilibrium between ionisation and recombination. In the present section we shall suppose that the ionisation takes place only under the action of the radiation from the central star. We shall later take into account the presence of the diffuse radiation of the nebula itself.

Let us find the number of ionisations occurring per unit volume and unit time. We may suppose that the ionisation in the nebulae takes place only from the ground state. Hence, denoting by n_1 the number of atoms in the ground state in unit volume, and by $k_{1\nu}$ the absorption coefficient referred to one atom, we obtain for the required number of ionisations

$$n_1 \int_{\nu_0}^{\infty} k_{1\nu} \frac{c\,\varrho_\nu}{h\,\nu}\, d\nu\,,$$

* A. YA. KIPPER, and L. SPITZER and J. L. GREENSTEIN [162], have recently put forward a new mechanism to explain the origin of the continuous spectrum of planetary nebulae, namely the emission of two photons by hydrogen atoms in passing from the $2\,s$ to the $1\,s$ state. Calculation shows that, of all the electrons captured into the second and higher levels, about 32 % should make the transition $2\,s \rightarrow 1\,s$ (if collisions are not important). The above authors have calculated the frequency distribution of the two-photon emission, and find that, in the visible region of the spectrum, the intensity varies relatively little. The intensity of the two-photon emission is comparable in order of magnitude with that due to recombinations and free-free transitions.

where ν_0 is the ionisation frequency and ϱ_ν the density of radiation in the nebula, equal to $W\varrho_\nu{}^*$.

Recombinations take place to all levels. Hence the total number of recombinations is

$$n_e\, n_+ \sum_1^\infty C_{if}(T_e)\ .$$

Since the number of ionisations must equal the number of recombinations, we have

$$n_1\, W \int_{\nu_0}^\infty k_{1\nu}\, \frac{c\,\varrho_\nu{}^*}{h\,\nu}\, \mathrm{d}\nu = n_e\, n_+ \sum_1^\infty C_{if}(T_e)\ . \qquad (24.20)$$

This is essentially the ionisation formula for the nebulae. However, it can be considerably simplified by using the relation which exists between the coefficients $k_{i\nu}$ and $C_{if}(T_e)$. The detailed derivation has been given in Chapter 5; here we shall recall the main ideas.

To derive the relation mentioned, we consider the state of thermodynamic equilibrium. In this case, as is well known, we have *detailed balancing*. In particular, the number of ionisations which take place from the ith level by the absorption of quanta with frequencies from ν to $\nu + \mathrm{d}\nu$ must equal the number of captures, into that level, of electrons with velocities from v to $v + \mathrm{d}v$, where

$$h\,\nu = \tfrac{1}{2}\, m\, v^2 + \chi_i\ .$$

We denote by $n_e\, f(v)\, \mathrm{d}v$ the number of free electrons with velocities between v and $v + \mathrm{d}v$ in 1 cm^3, and by $n_e\, n_+\, \beta_i(v)\, v\, f(v)\, \mathrm{d}v$ the number of captures of such electrons by ions, into the ith level, in 1 cm^3 in 1 second. From the above, we have

$$n_e\, n_+\, \beta_i(v)\, v\, f(v)\, \mathrm{d}v = 4\,\pi\, n_i\, k_{i\nu}(1 - e^{-h\nu/kT})\, I_\nu\, \mathrm{d}\nu/h\,\nu\ , \qquad (24.21)$$

where the factor $1 - e^{-h\nu/kT}$ takes account of negative absorption.

In thermodynamic equilibrium, however, the function $f(v)$ is determined by Maxwell's formula, the intensity of radiation I_ν by Planck's formula, and the distribution of atoms among the states by Boltzmann's and Saha's formulae. Using these formulae and (24.21), we obtain

$$\beta_i(v) = \frac{h^2\,\nu^2}{c^2\, m^2\, v^2}\, \frac{g_i}{g_+}\, k_{i\nu}\ , \qquad (24.22)$$

where g_i is the statistical weight of the ith state of the atom concerned, and g_+ that of the ground state of the ionised atom.

Formula (24.22) gives the required relation between the coefficients $\beta_i(v)$ and $k_{i\nu}$. Although it has been derived on the assumption of thermodynamic equilibrium, it is, of course, always valid.

It is evident that the coefficient $C_i(T_e)$ is [see (20.10)]

$$C_i(T_e) = \int_0^\infty \beta_i(v)\, v\, f(v)\, \mathrm{d}v\,. \tag{24.23}$$

Here $f(v)$ is given by the Maxwellian law of the velocity distribution of free electrons,

$$f(v) = \frac{4\,\pi\, m^3}{(2\,\pi\, m\, k\, T_e)^{3/2}}\, e^{-mv^2/2kT_e}\, v^2\,. \tag{24.24}$$

We must now substitute the value found for $C_i(T_e)$ in the relation (24.20). Before doing this, however, we rewrite it in the form

$$x\, n_1\, W \int_{\nu_0}^\infty k_{1\nu}\, \frac{c\, \varrho_\nu^*}{h\, \nu}\, \mathrm{d}\nu = n_e\, n_+\, C_1(T_e)\,, \tag{24.25}$$

where x is the proportion of captures into the first level. Substituting (24.23) in (24.25), we find

$$x\, n_1\, W \int_{\nu_0}^\infty k_{1\nu}\, \frac{\nu^2\, \mathrm{d}\nu}{e^{h\nu/kT_*} - 1} =$$

$$= \frac{g_1}{g_+}\, n_e\, n_+\, \frac{m\, h^3}{2\,(2\,\pi\, m\, kT_e)^{3/2}} \int_0^\infty k_{1\nu}\, \nu^2\, e^{-mv^2/2kT_e}\, v\, \mathrm{d}v\,. \tag{24.26}$$

The two integrals which appear in this relation are easily calculated if $k_{1\nu} \sim 1/\nu^2$. In actual fact, for hydrogen-like atoms $k_{1\nu} \sim 1/\nu^3$ (for other atoms the dependence of $k_{1\nu}$ on ν is imperfectly known). However, we shall nevertheless suppose that the absorption coefficient is inversely proportional to the square of the frequency, since the error thereby committed is very slight. As a result we obtain

$$n_e\, \frac{n_+}{n_1} = \frac{g_+}{g_1}\, x\, W\, \sqrt{\frac{T_e}{T_*}}\, \frac{2\,(2\,\pi\, m\, kT_*)^{3/2}}{h^3}\, \log_e\left(1 - e^{-h\nu_0/kT_*}\right)^{-1}\,. \tag{24.27}$$

Formula (24.27) can be somewhat simplified by noticing that the factors g_+/g_1, $2\,x$ and $\sqrt{(T_e/T_*)}$ do not differ greatly from unity, and moreover $h\,\nu_0 \gg k\,T_*$. Hence the ionisation formula for the nebulae takes the following approximate form:

$$n_e\, \frac{n_+}{n_1} = W\, \frac{(2\,\pi\, m\, kT_*)^{3/2}}{h^3}\, e^{-h\nu_0/kT_*}\,. \tag{24.28}$$

We see that formula (24.28) differs from the ordinary ionisation formula by the presence of the factor W on the right-hand side. This factor is very small for the nebulae ($W \approx 10^{-13}$). This however, does not mean that the degree of ionisation, i.e. the quantity n_+/n_1, is also small. In fact, the degree of ionisation in the nebulae may be very considerable, since the smallness of the dilution factor W is balanced by that of the concentration n_e of free electrons.

Strictly speaking, the approximate formula (24.28) is valid only when the optical thickness of the nebula beyond the limit of the principal series of the atom concerned is less than unity. Otherwise it is necessary to take account of the absorption of the radiation of the star, and also of the presence of the diffuse radiation of the nebula, which originates from recombinations to the first level. However, we shall show below that the diffuse radiation in the nebulae plays a very small part. It is therefore sufficient to take into account only the attenuation of the radiation which comes directly from the star. It is evident that this can be done by introducing the factor $e^{-\tau}$ on the right-hand side of formula (24.28), where τ is the optical distance from the nucleus beyond the principal series limit, which corresponds to some mean absorption coefficient. Thus we obtain, instead of formula (24.28),

$$n_e \frac{n_+}{n_1} = W \frac{(2\pi m k T_*)^{3/2}}{h^3} e^{-h\nu_0/kT_*} e^{-\tau}. \tag{24.29}$$

Let us assume that, for $\tau < 1$, the degree of ionisation is very high, i.e. $n_+/n_1 \gg 1$. When τ becomes of the order of unity, the degree of ionisation rapidly decreases. In turn, owing to the increase in the number of neutral atoms, this leads to a rapid increase in τ. Consequently, the transition from values of $\tau < 1$ to values of $\tau \gg 1$, and therefore the transition from $n_+/n_1 \gg 1$ to $n_+/n_1 \ll 1$, takes place in a comparatively short geometrical distance. Thus the nebula can be approximately divided into two regions with respect to the atoms or ions of a given kind: the *inner*, in which the degree of ionisation is determined by formula (24.28), and the *outer*, in which it is zero. The first region radiates in the lines of the atom concerned, while the second region does not. The boundary between these regions is where the optical distance from the nucleus, beyond the principal series limit, is of the order of unity.

It is clear that, in the case considered, different atoms will emit, generally speaking, in different volumes, i.e. there should be a "stratification" of the radiation in the nebulae. This is indeed the case: the images of the nebulae obtained by means of a slitless spectrograph are of different sizes in different lines. This indicates that, at least for some atoms, the optical thickness of the nebula, beyond the principal series

limit, exceeds unity. Hydrogen is perhaps an exception, since the dimensions of the images of the nebulae in the Balmer lines are usually the greatest, and consequently we have no certainty that beyond the regions of the nebulae which radiate in these lines there extend other regions which do not radiate in them.

Observation also reveals the following interesting fact. The higher the ionisation potential of the atom, the smaller the dimensions of the images of the nebulae in the lines of that atom. Thus, for instance, the dimensions of the images in the lines of ionised helium are considerably less than in those of neutral helium. This is easily explained. As an example, let us consider a helium nebula. If the temperature of the central star is fairly high, then, in the regions of the nebula which are closest to the star, there must be mainly singly and doubly ionised helium atoms, and so the radiation will be in the lines of He II. This region terminates where the optical distance from the nucleus, beyond the principal series limit of ionised helium, becomes of the order of unity. Beyond the boundaries of this region, the radiation from the star which is capable of doubly ionising helium does not penetrate, and only singly ionised and neutral atoms will be found there. This second region must radiate only in the lines of He I. It terminates, in turn, where the optical distance from the nucleus, beyond the principal series limit of neutral helium, becomes of the order of unity. The radiation from the star which is capable of ionising He I atoms will reach no further, and this outermost part of the nebula will not radiate at all.

The considerations given above are applicable not only to helium, but to other elements also. However, in actual nebulae, which consist of many elements, the picture is somewhat more complex, since the parts of the star's spectrum which are absorbed by various atoms and ions may overlap.

6. The masses of the nebulae. The results given above enable us to apply simple methods to estimate the masses and densities of gaseous nebulae. Turning to this question, let us assume that the most widely distributed element in the nebulae, as in stellar atmospheres, is hydrogen. In other words, we shall find the masses and concentrations of hydrogen in the nebulae.

The simplest, though a fairly crude, method of estimating the mass of a planetary nebula is based on the supposition that its optical thickness, beyond the limit of the Lyman series, is of the order of unity. The fact that this optical thickness cannot be very small compared with unity follows from the fact that otherwise the temperatures of the stars, as found by ZANSTRA's method, would have to be greatly increased. On the other hand, in the previous section we have expressed doubts that the

optical thickness concerned can considerably exceed unity. Thus we are justified in assuming that

$$\tau_0 = n_1 \, k \, r = 1 \; . \tag{24.30}$$

It is possible, of course, that our doubts were unfounded. If this is so, the equation (24.30) will refer only to the luminous part of the nebula. By using it we shall therefore obtain some lower limit to the required mass.

In the relation (24.30), k is the absorption coefficient beyond the limit of the Lyman series, referred to one atom. It is equal to $0{\cdot}5 \times 10^{-17}$. If we take for the radius of the nebula a value $r = 13{,}000$ astronomical units $= 2 \times 10^{17}$ cm, we obtain for the number of neutral hydrogen atoms in 1 cm^3 $n_1 = 1$.

In order to find the number of ionised hydrogen atoms, we use formula (24.28). For hydrogen this can be re-written in the form

$$n_e \, n_+/n_1 = 2{\cdot}44 \times 10^{15} \, W \, T_*^{3/2} \, e^{-157{,}200/T_*} \; . \tag{24.31}$$

For an average nebula we take $W = 10^{-14}$, $T_* = 40{,}000°$. We shall also assume $n_e = n_+$. With the value $n_1 = 1$ just found, formula (24.31) gives for the number of hydrogen ions in 1 cm^3 $n_+ = 2000$.

We see that in the planetary nebulae hydrogen is predominantly in the ionised state. Hence the mass of the nebula can be estimated from the formula

$$\mathsf{M} = \tfrac{4}{3} \, \pi \, r^3 \, n_+ \, m_{\mathrm{H}} \; , \tag{24.32}$$

where m_{H} is the mass of a hydrogen atom. This formula gives

$$\mathsf{M} = 10^{32} \, \mathrm{g} = 0{\cdot}05 \, \mathsf{M}_\odot \; .$$

Another, more exact, method of determining the masses of the nebulae is based on the use of formula (24.19), which determines the amount of energy emitted by the nebula in the hydrogen lines. This formula can be approximately written

$$E_{ki} = z_k \, A_{ki} \, h \, \nu_{ik} \, n_+^2 V \; . \tag{24.33}$$

From this we find for the concentration of hydrogen ions

$$n_+ = \sqrt{(E_{ki}/z_k A_{ki} h \nu_{ik} V)} \; , \tag{24.34}$$

and consequently for the mass of the nebula

$$\mathsf{M} = m_{\mathrm{H}} \sqrt{(E_{ki} \, V/z_k A_{ki} h \nu_{ik})} \; . \tag{24.35}$$

As before, if the optical thickness of the nebula, beyond the limit of the Lyman series, is large, formula (24.35) determines the mass not of the whole nebula, but only of its luminous part.

Let us pass, in formulae (24.34) and (24.35), from the energy E_{ki} to the total visual luminosity L of the nebula. We put

$$E_{ki} = \lambda_{ik} L .$$

Then we have instead of (24.34) and (24.35)

$$\varrho = m_{\mathrm{H}} n_{+} = C \sqrt{(L/V)} , \tag{24.36}$$

$$\mathsf{M} = C \sqrt{(L V)} , \tag{24.37}$$

where

$$C = m_{\mathrm{H}} \sqrt{(\lambda_{ik}/z_k A_{ki} h \nu_{ik})} . \tag{24.38}$$

Since the visual luminosity of the nebula is determined mainly by the N_1 and N_2 lines, the value of λ_{ik} is, roughly speaking, the ratio of the intensities of the given line and the $N_1 + N_2$ lines. For example, if the intensity ratio of the lines N_2/H_β is 3, the value of λ_{24} is 1/12. In this case $C = 1{\cdot}5 \times 10^{-12}$. Of course, the value of λ_{24} varies somewhat from one nebula to another. However, this has little effect on the value of C, since λ_{ik} appears under the radical sign in formula (24.38). Hence the factor C, to a first approximation, may be regarded as constant for all nebulae.

The calculation of the masses of planetary nebulae from formula (24.37) leads to values of the order of some hundredths of the Sun's mass, and the calculation of the concentrations from formula (24.36) gives values of the order of some thousands of atoms per cm^3. These values do not differ, on the average, from those obtained above, $\mathsf{M} = 0{\cdot}05\, \mathsf{M}_{\odot}$ and $n_+ = 2000$. The masses and densities of individual nebulae may apparently differ from the mean values mentioned by a factor of ten. This is chiefly caused by the dispersion of volumes, since the dispersion of the luminosities of the nebulae is very small.

The fact that the masses of the planetary nebulae amount to only a fraction of the Sun's mass again emphasises their subordinate position with respect to the nuclei. This is the more important because, according to modern ideas, the planetary nebulae were produced as a result of the ejection of matter from the central stars.

Formula (24.37) can also be applied to determine the masses of diffuse gaseous nebulae, which, as is well known, are luminous for the same reason as the planetary nebulae. However, in this case the volumes and luminosities (especially the latter) are still imperfectly known. Nevertheless, it may be asserted that the masses of the diffuse nebulae may reach hundreds or thousands of times the mass of the Sun. This fact also is undoubtedly of great cosmogonical significance.

7. **The chemical composition of the nebulae.** In the preceding section we have determined the concentration of hydrogen atoms in the planetary nebulae. We shall now show how the concentration of other atoms (relative to hydrogen) may be found. Here we shall suppose known the ratios, obtained from observation, of the intensities of the lines of the given element to those of the lines of the Balmer series.

Let us first assume that the lines of the (neutral or ionised) atom considered are formed as a result of recombinations. We construct, for each level of this atom, the equations of the steady state, similar to the equations (24.17) for hydrogen. This system of equations gives us the quantities $z_m' = n_m'/n_e n_+'$, where n_m' is the number of atoms in the mth state and n_+' is the number of atoms in the next ionisation state (per unit volume). Using the values of z_m', we obtain for the energy emitted by the nebula in that line of the atom concerned which is formed by the transition $m \to l$

$$E_{ml}' = z_m' A_{ml}' h\nu_{lm}' n_e n_+' V' ,$$

where V' is the volume which radiates in the line considered. We write the similar expression for the energy emitted by the nebula in the Balmer line corresponding to the transition $k \to 2$:

$$E_{k2} = z_k A_{k2} h\nu_{2k} n_e n_+ V .$$

From these two formulae we have

$$\frac{E_{ml}'}{E_{k2}} = \frac{z_m' A_{ml}' \nu_{lm}'}{z_k A_{k2} \nu_{2k}} \frac{n_+'}{n_+} \frac{V'}{V} . \tag{24.39}$$

Since the ratio of intensities of the lines E_{ml}'/E_{k2} can be obtained from observation, formula (24.39) makes it possible to determine n_+'/n_+, i. e. the ratio of the number of atoms of the given element in some ionisation state to the number of hydrogen ions.

In order to obtain the number of atoms in other ionisation states, the ionisation formula (24.28) must be used. We then find the ionisation state in which are the majority of the atoms of the element concerned. The total concentration of this element in the nebula will thus be determined.

The calculations described are most easily performed for ionised helium (and other hydrogen-like ions), since in this case the quantities z_m' can be obtained from the z_m for hydrogen. Thus, if Z is the atomic number of the hydrogen-like ion, its energy levels lie Z^2 times as deep as the hydrogen levels, and the expressions which determine the probabilities of recombinations and of spontaneous transitions are obtained

from the corresponding expressions for hydrogen by multiplying by Z^4. For this reason, we find, by considering equations (24.17) and formula (24.14), that the numbers z_m' for ionised helium are one-eighth of the numbers z_m for hydrogen, if the temperature is taken at four times the value, i. e.

$$z_m'(4T_e) = \tfrac{1}{8} z_m(T_e) .$$

The calculation of the quantities z_m' for neutral helium has been carried out by A. A. Nikitin [114]. The relative intensities of the helium lines which he obtained were in satisfactory agreement with observation for both the singlet and the triplet series. Using formula (24.39), he found that the number of helium atoms in the nebulae is, on the average, one-tenth of the number of hydrogen atoms.

The determination of the concentration of other atoms in the nebulae offers considerable difficulties, since the probabilities of spontaneous transitions and of capture are very imperfectly known for them. Furthermore, there is some uncertainty connected with the transition from one ionisation state to another by using the ionisation formula, because of the probable deviation of the radiation of the stars from Planck's Law in the far ultra-violet region of the spectrum. However, approximate estimates are possible in these cases.

Another possibility for determining the concentration of atoms in the nebulae exists by reason of the presence of forbidden lines in their spectra which are excited by electron collisions. Let us find, as an example, the concentration of doubly ionised oxygen atoms, using the observed ratio of intensities of the lines $N_1 + N_2$ and H_β.

We denote by $n_{\mathrm{O\,III}}$ the number of O III ions in 1 cm^3. These ions are predominantly in the first state. In collisions with free electrons, transitions take place from the first state to the second. The number of such transitions occurring in 1 cm^3 in 1 second is written in the form $n_{\mathrm{O\,III}}\, n_e\, b_{12}'(T_e)$. Almost all the atoms which have gone from the first state to the second later return spontaneously, emitting quanta in the N_1 and N_2 lines. Hence the total energy radiated by the nebula in the N_1 and N_2 lines is

$$E_{N_1+N_2} = n_{\mathrm{O\,III}}\, n_e\, b_{12}'(T_e)\, h\nu_{12}'\, V' ,$$

where V' is the volume of the nebula which radiates in these lines. On the other hand, we have for the energy radiated by the nebula in the H_β line

$$E_{H_\beta} = z_4\, A_{42}\, h\nu_{24}\, n_e\, n_+\, V .$$

The last two formulae give

$$\frac{E_{N_1+N_2}}{E_{H_\beta}} = \frac{4}{3}\frac{E_{N_1}}{E_{H_\beta}} = \frac{n_{\mathrm{O\,III}}}{n_+}\,\frac{b_{12}'(T_e)}{z_4 A_{42}}\,\frac{\nu_{12}'}{\nu_{24}}\,\frac{V'}{V} . \qquad (24.40)$$

To find the ratio $n_{\mathrm{O\,III}}/n_+$ by means of this formula, it is necessary to know the ratios E_{N_2}/E_{H_β} and V'/V from observation. In the spectra of planetary nebulae, the N_2 line is, on the average, three times as bright as the H_β line, and the images of the nebulae obtained by means of the slitless spectrograph are approximately the same size in these lines. We therefore take $E_{N_1+N_2}/E_{H_\beta} = 12$, $V'/V = 1$.

The quantity $b_{12}'(T_e)$ depends very strongly on T_e. However, the electron temperatures of the nebulae are known to within fairly narrow limits. We shall take $T_e = 8000°$. For this temperature, $b_{12}' = 0{\cdot}5 \times 10^{-8}$. Next, from Table 18 we find, for the same temperature, $z_4 = 0{\cdot}4 \times 10^{-20}$. Finally, we have $A_{42} = 8{\cdot}37 \times 10^6$ and $\nu_{12}'/\nu_{24} \approx 1$ (since the H_β and N_2 lines lie close together). Using these data, we obtain from fromula (24.40) $n_{\mathrm{O\,III}}/n_+ = 10^{-4}$. Thus, in the planetary nebulae there are ten thousand hydrogen ions to every O III ion.

It is important to notice that O II lines arising as a result of recombinations are observed in the spectra of the nebulae. This makes it possible to estimate the number of O III ions by the first of the methods described above also. Good agreement is obtained between the two estimates, which is an indication of their correctness. A similar test of the values found for the concentrations can be made for some other atoms also.

The concentration of various atoms in the planetary nebulae has been determined by a number of authors, using the methods explained above. The results are given in Table 1 (see Section 5.1).

Taking into consideration the inexactness of such determinations, we may conclude that there are no great differences in chemical composition between the nebulae and stellar atmospheres. This conclusion can be formulated as follows: the outer layers of the Sun, if expanded to the dimensions of a nebula and irradiated by a hot star, should give exactly the same spectrum as is observed for the planetary nebulae.

Chapter 25. Radiative equilibrium in planetary nebulae

1. The Lyman continuum radiation field. In determining the intensities of emission lines, we have assumed that the nebulae are transparent to radiation in these lines. Such an assumption is not suspect as regards the lines of subordinate series, since only a negligible minority of the atoms are in the excited states. It is justified also as regards forbidden lines (even if the lower state is the ground state), since the absorption coefficient in these lines is extremely small.

However, the nebulae are not in general transparent to radiation in the frequencies of a principal series. This greatly complicates the calculation of the radiation field in these frequencies, since it is then necessary to use the equations of radiative transfer. We shall now calculate the radiation field in the frequencies of the Lyman series of hydrogen. This problem was first considered by V. A. AMBARTSUMYAN [5]. Later, a number of authors refined and generalised the solution which he obtained.

In investigating radiative transfer in the nebulae, the geometrical model of the nebula is of great importance. We shall assume (as is usually done) that the nebula is bounded by two concentric spheres of radii r_1 and r_2. The nucleus of the nebula is at the centre of these spheres. The thickness of the nebula is assumed small in comparison with its distance from the nucleus (i. e. $r_2 - r_1 \ll r_1$). In this case the nebula can be regarded as consisting of plane-parallel layers, and the dilution factor is constant in the nebula.

We first examine the radiation field in the Lyman continuum. When the L_c quanta which come from the star to the nebula are absorbed, hydrogen atoms are ionised. In recombinations to the first level, the quanta are emitted again. We denote by x the proportion of recombinations which are to the first level. We can then say that a scattering of L_c quanta takes place in the nebulae, the probability of "survival" of the quanta at each elementary scattering process being x.

For simplicity, we shall consider the entire Lyman continuum as a single level. Let $\varkappa$ be the mean atomic absorption coefficient in the continuum and τ the corresponding optical depth, measured from the inner boundary of the nebula, i. e.

$$\tau = \int_{r_1}^{r} n_1 \varkappa \, \mathrm{d}r \, .$$

We denote by πS the number of quanta in the Lyman continuum incident from the star on 1 cm^2 of the inner boundary of the nebula. It is evident that, of these, $n_1 \varkappa \pi S e^{-\tau}$ quanta are absorbed in 1 cm^3 at optical depth τ. However, besides the radiation which comes directly from the star, there is absorbed in this volume also the diffuse radiation of the nebula itself, originating in recombinations to the first level. The number of quanta of diffuse L_c radiation absorbed in 1 cm^3 is $n_1 \varkappa \int K(\tau, \theta) \, \mathrm{d}\omega$, where $K(\tau, \theta) \, h\nu_0$ is the intensity of radiation at an angle θ to the normal at optical depth τ, and the integration is taken over the whole solid angle.

Of the total number of quanta

$$n_1 \varkappa \pi S e^{-\tau} + n_1 \varkappa \int K(\tau, \theta) \, \mathrm{d}\omega$$

which are absorbed in unit volume, a fraction x passes again into L_c radiation. Hence, denoting by $4\pi n_1 \varkappa C(\tau)$ the number of quanta emitted by this volume in the Lyman continuum, we obtain

$$C(\tau) = x \int K(\tau,\theta) \frac{d\omega}{4\pi} + x \frac{S}{4} e^{-\tau} . \tag{25.1}$$

This is the condition of radiative equilibrium for L_c radiation in a nebula.

As well as by equation (25.1), the quantities $C(\tau)$ and $K(\tau,\theta)$ are related by the equation of radiative transfer, which, in the case of plane-parallel layers, has the form

$$\cos\theta \, dK(\tau,\theta)/d\tau = C(\tau) - K(\tau,\theta) . \tag{25.2}$$

The equations (25.1) and (25.2) are to be solved with the following boundary conditions:

$$\left.\begin{aligned} K(0,\theta) &= K(0, \pi - \theta) , \\ K(\tau_0,\theta) &= 0 \text{ for } \theta > \tfrac{1}{2}\pi . \end{aligned}\right\} \tag{25.3}$$

The first of these conditions, which holds for the inner boundary of the nebula (where $\tau = 0$), means that the intensity of the radiation emerging from the nebula is equal to that of the radiation entering it. This is because the radiation leaving the nebula at any point on the inner boundary at an angle θ to the normal is also the radiation which enters the nebula at an angle $\pi - \theta$ (Fig. 61). The second condition expresses the obvious fact that, at the outer boundary of the nebula (where $\tau = \tau_0$), there is no incident radiation.

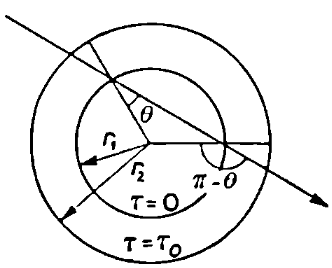

Fig. 61

Thus our problem consists in solving the equations (25.1) and (25.2) with the boundary conditions (25.3). These equations can easily be solved by one of the approximate methods which are known in the theory of radiative equilibrium (by averaging the intensity of radiation with respect to direction). However, we shall prefer to construct an integral equation. Solving equation (25.2) for $K(\tau,\theta)$ with the conditions (25.3), and substituting the expression found for $K(\tau,\theta)$ in equation (25.1), we obtain

$$C(\tau) = \tfrac{1}{2} x \int_0^{\tau_0} [E_1(|\tau - \tau'|) + E_1(\tau + \tau')] \, C(\tau') \, d\tau' + \tfrac{1}{4} x S e^{-\tau} , \tag{25.4}$$

where

$$E_1(\tau) = \int_{\tau}^{\infty} e^{-t} \frac{dt}{t} .$$

It is clear that, the greater the optical thickness of the nebula, the greater the part played by the diffuse radiation in it. Since, however, our main purpose is to ascertain the importance of the diffuse radiation, we shall put $\tau_0 = \infty$. In this case, equation (25.4) can be approximately rewritten in the form

$$C(\tau) = \tfrac{1}{2}\, x \int_{-\infty}^{\infty} E_1(|\,\tau - \tau'\,|)\, C(\tau')\, d\tau' , \tag{25.5}$$

and equation (25.5) determines the function $C(\tau)$ the more accurately, the greater the optical depth τ.

Equation (25.5), however, has the exact solution

$$C(\tau) = Ae^{-k\tau} , \tag{25.6}$$

where k satisfies the equation

$$\frac{x}{2k} \log_e \frac{1+k}{1-k} = 1 , \tag{25.7}$$

and A is an arbitrary constant. We shall regard the function (25.6) as an approximate solution of equation (25.4), and find the constant A from the condition that this equation is exactly satisfied on the average. We then obtain

$$A = k\, x\, S/4\,(1 - x) . \tag{25.8}$$

The values of k found from equation (25.7) are given below.

x	0	0·5	0·6	0·7	0·8	0·9	1·0
k	1·00	0·96	0·91	0·83	0·71	0·53	0

For L_c radiation the value of x is

$$x = C_1(T_e) / \sum_{i=1}^{\infty} C_i(T_e) ,$$

where $C_i(T_e)$ is determined by formula (24.14). Calculations from this formula give

T_e	5000°	10,000°	20,000°	50,000°
x	0·39	0·44	0·49	0·57

Since the electron temperatures of the nebulae are close to 10,000°, we obtain from the tables given above $k = 0{\cdot}97$. We shall take simply $k = 1$. In this case we find for the function $C(\tau)$

$$C(\tau) = \frac{x\,S}{4\,(1-x)}\,e^{-\tau}\,. \tag{25.9}$$

On comparing the expression obtained for $C(\tau)$ with (25.1), we see that the number of quanta of diffuse L_c radiation is approximately equal to the number of L_c quanta coming directly from the star. Thus it must be acknowledged that the diffuse radiation plays a fairly small part.

This result is explained by the fact that the proportion of captures which are into the first level, i. e. the value of x, is comparatively small (less than $\frac{1}{2}$). If x were close to unity, the diffuse radiation would predominate over the direct radiation. This would be particularly noticeable at large optical depths (in consequence of the small value of k in this case).

2. **The Lyman α radiation field.** Much more interesting results than in the preceding case are obtained from an analysis of the Lyman α radiation field in the nebulae. We have seen above that, for every L_c quantum that is transformed in the nebula, one Lyman α quantum is necessarily formed. These quanta then diffuse in the nebula, undergoing pure scattering. However, the optical thickness of the nebula in the Lyman α line is 10^4 to 10^5 times that in the Lyman continuum, i. e. it is very large. This has the result that each Lyman α quantum, before leaving the nebula, must undergo a very large number of scattering processes. Consequently it can be predicted that the density of Lyman α radiation in the nebulae should be extremely high.

Let s_{12} be the absorption coefficient in the Lyman α line, referred to one atom. We introduce the optical depth t and the optical thickness t_0 in this line,

$$t = \int_{r_1}^{r} n_1\, s_{12}\, \mathrm{d}r\,,\quad t_0 = \int_{r_1}^{r_2} n_1\, s_{12}\, \mathrm{d}r\,. \tag{25.10}$$

We write

$$t/\tau = t_0/\tau_0 = s_{12}/\varkappa = 1/q\,. \tag{25.11}$$

As we have already said, this ratio is of the order of 10^4 to 10^5.

We denote by $K_{12}(t,\theta)\,h\nu_{12}$ the intensity of Lyman α radiation at an angle θ to the normal at optical depth t, and by $4\,\pi\,n_1\,s_{12}\,C_{12}(t)$ the number of Lyman α quanta emitted in unit time and unit volume at that depth. The quantities $K_{12}(t,\theta)$ and $C_{12}(t)$ are related by the usual equation of radiative transfer

$$\cos\theta\,\mathrm{d}K_{12}(t,\theta)/\mathrm{d}t = C_{12}(t) - K_{12}(t,\theta)\,. \tag{25.12}$$

Let us now construct the equation of radiative equilibrium for the Lyman α radiation. Here we shall assume that all the Lyman α quanta are formed from L_c quanta, i. e. we neglect the radiation of the star in the Lyman lines (including the Lyman α line itself) in comparison with its radiation in the Lyman continuum. It is clear that this assumption is fully justified so long as τ_0 is not very small compared with unity.

In the previous section we have seen that, as a result of the absorption of L_c radiation, $4\pi n_1 \varkappa C(\tau)$ recombinations to the first level take place in unit volume per unit time. It is evident that the number of recombinations to all other levels is obtained from this by multiplying by $(1-x)/x$. But each such recombination necessarily leads to the formation of a Lyman α quantum. Hence the number of Lyman α quanta originating from L_c radiation is

$$4\pi \frac{1-x}{x} n_1 \varkappa C(\tau) .$$

However, the volume element does not emit only these quanta in the Lyman α line. It scatters a much larger number of quanta. More exactly, it emits as many quanta as it absorbs from the diffuse Lyman α radiation field. (It is clear that this diffuse Lyman α radiation has also been formed from L_c radiation, but has not yet succeeded in escaping from the nebula.) Thus the total number of Lyman α quanta emitted by unit volume is

$$n_1 s_{12} \int K_{12}(t,0)\,\mathrm{d}\omega + 4\pi \frac{1-x}{x} n_1 \varkappa C(\tau) .$$

This number of quanta was denoted above by $4\pi n_1 s_{12} C_{12}(t)$. We therefore obtain, as the condition of radiative equilibrium,

$$C_{12}(t) = \int K_{12}(t,0) \frac{\mathrm{d}\omega}{4\pi} + q \frac{1-x}{x} C(\tau) . \tag{25.13}$$

The function $C(\tau)$ which appears in this equation is given by formula (25.9). Using this, we finally have

$$C_{12}(t) = \int K_{12}(t,0) \frac{\mathrm{d}\omega}{4\pi} + q \frac{S}{4} e^{-qt} . \tag{25.14}$$

Let us solve equations (25.12) and (25.14) in Eddington's approximation. Let $\bar{K}_{12}(t)\, h\nu_{12}$ be the mean intensity of Lyman α radiation, and $4\pi \bar{H}_{12}(t)\, h\nu_{12}$ the flux of this radiation, i. e.

$$\bar{K}_{12}(t) = \int K_{12}(t,0) \frac{\mathrm{d}\omega}{4\pi} , \quad \bar{H}_{12}(t) = \int K_{12}(t,0) \cos\theta \frac{\mathrm{d}\omega}{4\pi} . \tag{25.15}$$

The equations mentioned give

$$\left.\begin{aligned} \bar{K}_{12}(t) &= a - 3\,bt - (3\,S/4\,q)\,e^{-qt}\,, \\ \bar{H}_{12}(t) &= b - \tfrac{1}{4}\,Se^{-qt}\,, \end{aligned}\right\} \tag{25.16}$$

where a and b are arbitrary constants.

To find these arbitrary constants, it is necessary to specify the boundary conditions. Here two cases must be distinguished.

(1) The nebula is stationary (or is expanding with a velocity not exceeding the mean thermal velocity of the atoms). In this case the boundary conditions are the same as the conditions (25.3) for L_c radiation. In the approximation considered, they may be written

$$\bar{H}_{12}(0) = 0,\; 2\,\bar{H}_{12}(t_0) = \bar{K}_{12}(t_0)\,. \tag{25.17}$$

(2) The velocity of expansion of the nebula is large compared with the mean thermal velocity of the atoms. In this case the Lyman α quanta coming from one side of the nebula will not be absorbed by atoms on the opposite side, because of the Doppler effect. Hence the condition at the inner boundary of the nebula becomes similar to that at the outer boundary, i. e. we have

$$2\,\bar{H}_{12}(0) = -\bar{K}_{12}(0),\; 2\,\bar{H}_{12}(t_0) = \bar{K}_{12}(t_0)\,. \tag{25.18}$$

We shall assume for simplicity that the optical thickness of the nebula in the Lyman continuum is considerably greater than unity, i. e. $qt_0 \gg 1$. Then we obtain for the constants a and b, in the first case mentioned,

$$a = \tfrac{3}{4}\,St_0,\; b = \tfrac{1}{4}\,S\,, \tag{25.19}$$

and in the second case,

$$a = \tfrac{3}{4}\,S/q,\; b = \tfrac{1}{4}\,S/qt_0\,. \tag{25.20}$$

By substituting the values found for a and b in formulae (25.16) it is possible to calculate the density and flux of Lyman α radiation at any optical depth in the nebula. Let us find, for example, the density of Lyman α radiation at the inner boundary of the nebula, in the first case considered above. It is evident that the number of Lyman α quanta in 1 cm³ is $4\,\pi\,\bar{K}_{12}(t)/c$. For $t = 0$ we obtain

$$\frac{4\,\pi}{c}\,\bar{K}_{12}(0) = \frac{3\,\pi\,S}{c}\,t_0\,.$$

But $\pi S/c$ is the number of L_c quanta in 1 cm³ at $t = 0$. Consequently, the density of Lyman α radiation at the inner boundary of the nebula is $3\,t_0$ times, i. e. some tens of thousands of times, the density of L_c radiation. In turn, the density of L_c radiation for a temperature of the star of the order of 40,000° to 50,000° is approximately 5×10^4 times the density of Lyman α radiation coming from the star. Hence the density of diffuse Lyman α radiation at the inner boundary of the nebula is 10^9 times what it would be if there were only the direct Lyman α radiation of the nucleus. This density is very much greater than the density of radiation in all the other lines of the hydrogen spectrum. However, if W is of the order of 10^{-13}, the density is still 10^4 times less than the density of Lyman α radiation at the surface of the star.

The two following questions are of interest in connection with such a high density of Lyman α radiation in the nebula: (1) what is the mean time during which a Lyman α quantum remains in the nebula and (2) what is the mean number of scatterings undergone by a Lyman α quantum?

In order to answer the first of these questions, we must know the total number of Lyman α quanta in the nebula, and the number originating from L_c radiation (or leaving the nebula) per unit time. It is evident that the ratio of these two quantities gives us the required time during which a quantum remains in the nebula. As we have already remarked, there are $4\pi \bar{K}_{12}(t)/c$ Lyman α quanta in 1 cm³. Hence the total number of Lyman α quanta in the nebula is

$$4\pi r_1^2 \frac{4\pi}{c} \int_0^{t_0} \bar{K}_{12}(t) \frac{dt}{n_1 s_{12}} .$$

The number of Lyman α quanta originating in the nebula is simply equal to the number of L_c quanta arriving from the central star. Thus $4\pi r_1^2 \pi S$ Lyman α quanta are formed in the nebula in 1 second. Consequently, we obtain for the mean time during which a quantum remains in the nebula

$$T = \frac{4}{cS} \int_0^{t_0} \bar{K}_{12}(t) \frac{dt}{n_1 s_{12}} . \tag{25.21}$$

To calculate the integral, we need an expression for $\bar{K}_{12}(t)$. Substituting the constants a and b, determined by formula (25.19), in the first of formulae (25.16), and omitting the last term (which is small by virtue of our assumption $qt_0 \gg 1$), we find

$$\bar{K}_{12}(t) = \tfrac{3}{4} S(t_0 - t) . \tag{25.22}$$

For simplicity we shall assume that $n_1 \varkappa =$ constant. The integration in formula (25.21) then gives

$$T = 3\, t_0^2/2\, c\, n_1\, s_{12}\,. \tag{25.23}$$

Having found the value of T, there is no difficulty in determining the mean number of scatterings undergone by a Lyman α quantum. It is evident that, to determine this number, we must divide T by the mean time interval between two successive scatterings. The mean path between two scatterings is $1/n_1\, s_{12}$ (corresponding to unit optical distance). The mean time interval between scatterings is therefore $1/c\, n_1\, s_{12}$. Dividing T by $1/c\, n_1\, s_{12}$, we obtain

$$N = \tfrac{3}{2}\, t_0^2\,. \tag{25.24}$$

Thus the mean number of scatterings of a Lyman α quantum is equal, in order of magnitude, to the square of the optical thickness of the nebula in the Lyman α line.

Let us find the numerical values of N and T. Since the optical thickness of the nebula for Lyman α radiation is 10^4 to 10^5, we find from formula (25.24) that each Lyman α quantum undergoes on the average some thousand million scatterings. To estimate the value of T from formula (25.23), we put $c = 3 \times 10^{10}$ cm/sec, $n_1 = 1$, $s_{12} = 10^{-12}$ cm^2. With these values, the mean time during which a Lyman α quantum remains in the nebula is found to be of the order of a thousand years.

The values obtained are so large that they compel us to examine closely all the agencies whereby the process of multiple scattering of the Lyman α quanta may be curtailed. One of these is collisions of the second kind, which transfer the hydrogen atoms from the second state to the first without the emission of a Lyman α quantum. Let us consider whether collisions of the second kind can noticeably diminish the density of Lyman α radiation in the nebula.

We have just seen that each Lyman α quantum undergoes 10^9 scatterings, i. e. brings atoms 10^9 times from the ground state to the excited state. The lifetime of the atom in the excited state is 10^{-8} sec. Consequently, the Lyman α quantum is in the absorbed state for 10 sec in all. The number of collisions of the second kind which occur in this time is

$$10\, n_e\, v_e\, \sigma\,,$$

where n_e is the number of free electrons in 1 cm^3, v_e is the mean velocity of the electrons, and σ is the effective cross-section for the process considered. The most probable values of these quantities are $n_e = 10^4$, $v_e = 10^8$ cm/sec and $\sigma = 10^{-16}$ cm^2. Hence the required number of

collisions of the second kind is $10\, n_e\, v_e\, \sigma = 10^{-3}$. However we alter the original data, this figure cannot approach unity. Consequently, the great number of scatterings of Lyman α quanta will hardly ever be interrupted as a result of collisions of the second kind.

Thus collisions cannot significantly diminish the density of Lyman α radiation. However, the results obtained in this section are not entirely applicable to actual nebulae. The reason for a considerable decrease in the density of Lyman α radiation in the nebulae will be demonstrated in the next section.

3. **Radiation pressure in the nebulae.** Let us now consider the flux of Lyman α quanta. At an optical depth t, the flux is $4\pi\, \bar{H}_{12}(t)$. Let us find the value of the flux at the boundaries of the nebula. To do so, we must substitute the values of the constants a and b in the second of formulae (25.16) and put in it $t = 0$ and $t = t_0$. If the nebula is stationary, the flux is zero at the inner boundary and $+\pi S$ at the outer boundary. If the nebula is expanding with a high velocity, the flux is $-\pi S\,[1-(1/qt_0)]$ at the inner boundary and $+\pi S/qt_0$ at the outer boundary. Since we suppose that $qt_0 \gg 1$, in the second case nearly all the Lyman α quanta emerge from the nebula at its inner boundary, and the fraction of quanta that do so is the greater, the greater the optical thickness of the nebula. In both cases, as we should expect, the total number of Lyman α quanta emerging from the nebula is $4\,\pi\, r_1^2\, \pi\, S$, which is the number of L_c quanta coming from the star.

Thus the flux of Lyman α quanta in the nebulae is very large. We should therefore expect that the radiation pressure due to the Lyman α radiation is also large. It will at any rate be much greater than that due to the L_c radiation, since the absorption coefficient in the Lyman α line is tens of thousands of times greater than that in the Lyman continuum.

Let us find the ratio of the radiation pressure caused by the Lyman α radiation to the force of gravity from the nucleus. We take a unit volume at one of the boundaries of the nebula (the outer boundary in the case of a stationary nebula and the inner boundary for an expanding one). The radiation pressure in the Lyman α line acting on this volume is

$$R = \frac{n_1 s_{12}}{c}\, \pi\, S\, h\nu_{12}\,. \tag{25.25}$$

From the definition of $\pi\, S$,

$$4\,\pi\, r_1^2\, \pi\, S = 4\,\pi\, r_*^2\, \frac{2\,\pi}{c^2} \int\limits_{\nu_1}^{\infty} \frac{\nu^2\, d\nu}{e^{h\nu/kT_*} - 1}\,. \tag{25.26}$$

We therefore obtain

$$R = \left(\frac{r_*}{r_1}\right)^2 \frac{2\pi n_1 s_{12} h\nu_{12}}{c^3} \int_{\nu_0}^{\infty} \frac{\nu^2 \, d\nu}{e^{h\nu/kT_*} - 1} \, . \tag{25.27}$$

To find the force of gravity from the nucleus acting on the same volume, it is necessary to take into account the fact that there are ionised atoms of hydrogen as well as neutral ones. The force of gravity is therefore

$$G = \left(\frac{r_*}{r_1}\right)^2 g_* \, m_{\mathrm{H}} \, (n_1 + n_+) \, , \tag{25.28}$$

where g_* is the acceleration due to gravity at the surface of the star.

For the required ratio R/G, we find

$$\frac{R}{G} = \frac{2\pi s_{12} h\nu_{12}}{c^3 m_{\mathrm{H}} g_* (1 + n_+/n_1)} \int_{\nu_0}^{\infty} \frac{\nu^2 \, d\nu}{e^{h\nu/kT_*} - 1} \, . \tag{25.29}$$

We put $T_* = 40{,}000°$, $n_+/n_1 = 5000$. Then formula (25.29) gives

$$R/G = 10^9/g_* \, . \tag{25.30}$$

It is difficult to suppose that the acceleration due to gravity at the surface of the central star is as much as 10^9 cm/sec^2. In fact, if this were so, the mass of the central star would be several thousand times that of the Sun. In that case, however, we should observe a red-shift of the absorption lines in the spectra of the nuclei corresponding to velocities of the order of 10,000 km/sec. Observations seem to indicate such a red-shift in the spectra of the nuclei, but it does not exceed 100 km/sec. Hence the masses of the nuclei cannot in reality exceed that of the Sun by a factor of more than ten.

Thus we reach the conclusion that the radiation pressure due to the Lyman α radiation plays a much greater part in the nebulae than does the force of gravity from the nucleus. The radiation pressure must be especially great at the boundaries of the nebulae: at the outer boundary of a stationary nebula, where it is directed outwards, and at the inner boundary of an expanding nebula, where it is directed towards the star. In the latter case, the radiation pressure must retard the inner parts of the nebula. The magnitude of this effect may be calculated. It is found that the retardation amounts to 3 km/sec per century.

The results obtained relate, however, only to a stationary nebula or to one expanding without a velocity gradient. But even if this were so at some moment, a difference in the expansion velocity would inevit-

ably occur by degrees, owing to radiation pressure. It is therefore necessary to consider the Lyman α radiation field in a nebula expanding with a velocity gradient.

It is clear that the appearance of a velocity gradient in the nebula should diminish the density of Lyman α radiation. This is due to the fact that, for a nebula expanding with a velocity gradient, the radiation quanta are able to emerge not only from the boundary regions, but also from the interior regions as a result of the Doppler effect. Thus, let us assume that the outer parts of the nebula are expanding more rapidly than the inner ones. We consider an atom in the central parts of the nebula, having a large thermal velocity component perpendicular to the layers. It is evident that the radiation quanta emitted by this atom in the direction opposite to its motion will hardly be absorbed at all in the nebula, whilst the atom will absorb from all directions approximately as many quanta as do the other atoms. Hence the Lyman α quanta, wherever they are, have a considerable probability of emerging from the nebula as a result of scattering by rapidly moving atoms.

A detailed analysis of the Lyman α radiation field in a nebula expanding with a velocity gradient has been carried out by V. V. Sobolev [153]. He has shown that the appearance of even a small velocity gradient leads to a very great diminution in the density and flux of Lyman α radiation, and therefore in the radiation pressure due to this radiation. It must be supposed that in actual nebulae the radiation pressure due to Lyman α radiation is comparable, in order of magnitude, with the force due to gravity from the central star.

It should be mentioned that V. V. Sobolev's work was based on the hypothesis of the complete frequency redistribution of the radiation in each elementary scattering process (i. e. the hypothesis that scattering is completely non-coherent). Subsequently, H. Zanstra [183] discussed, on the same hypothesis, the problem of radiation pressure in a stationary nebula. He finds that, in a nebula whose optical thickness in the Lyman continuum is of the order of unity, the force of radiation pressure for complete frequency redistribution is about 300 times less than the force when there is no frequency redistribution. V. V. Sobolev [161] has recently considered the problem of radiation pressure in a nebula for three cases of diffusion of radiation: (1) without change of frequency, (2) with complete frequency redistribution, (3) with the actual frequency redistribution. The latter case is taken to be the diffusion of radiation with the frequency redistribution caused by the natural width of the atomic energy levels and by the Doppler effect from the thermal motion of the atoms. It was found that cases 2 and 3 are fairly similar, but are very different from case 1. This gives some support for the hypothesis of complete frequency redistribution that was used in previous work.

4. **The problem of the origin of the planetary nebulae.** The results, given above, of the study of planetary nebulae compel us to suppose that there is a genetic relation between the nebula and its nucleus. It is most natural to assume that the nebulae are the remains of matter ejected at some time from the central star. This hypothesis is supported, first of all, by the expansion of the planetary nebulae. Another confirmation of this assumption is given by the results of the determination of the masses of the nebulae. It is clear that, if the masses of the nebulae were found to be greater than the masses of the nuclei, this hypothesis would become improbable. However, the masses of the nebulae, as we know, do not in reality exceed one-tenth of the Sun's mass.

It was thought not long ago that the planetary nebulae are formed in the eruptions of supernovae. As we shall see below, a mass of the order of the Sun's mass is ejected in such a major catastrophe, and this is quite sufficient for the formation of a planetary nebula. However, this hypothesis must now be abandoned. One of the objections which may be raised against it appears when we consider the expansion velocities of the nebulae. It is known that these velocities are very small, about 10 to 20 km/sec, whereas in the eruptions of supernovae the velocities with which matter is ejected are some thousands of kilometres per second. It is impossible to explain this difference by a single gravitational retardation, and we can point to no other retarding forces. In particular, they cannot be the forces of radiation pressure in the Lyman α line, since, as was established above, a velocity gradient arises as a result of the action of these forces, and the radiation pressure is then reduced.

The results of observations of the Crab Nebula also contradict the above hypothesis. It can scarcely be doubted that this nebula was formed by the eruption of the supernova of 1054. However, this nebula is at the present time, i. e. 900 years after the eruption, expanding with a velocity of 1300 km/sec, and it can be shown that it has not undergone any considerable retardation. Furthermore, in its outward appearance the Crab Nebula is not at all similar to the planetary nebulae.

Finally, the hypothesis of the origin of planetary nebulae in supernova eruptions is also contradicted by statistical considerations. Statistics show that eruptions of supernovae occur too rarely to bring about the observed number of planetary nebulae. We see more than 300 nebulae lying comparatively close to the Sun. Extrapolating, we find that there are approximately 10,000 nebulae in the entire Galaxy. The radii of the nebulae do not exceed 1 parsec, i. e. 3×10^{18} cm. This means that at larger dimensions the nebulae cease to be visible. Let a nebula be expanding with a velocity of 10 km/sec, so that its radius increases by 3×10^{13} cm per year. Then the lifetime of a nebula is 10^5 years.

It is evident that if a nebula exists for 10^5 years, and the total number of nebulae in the Galaxy is 10^4, then a nebula must appear every ten years. However, eruptions of supernovae take place much more rarely than once a decade. In our Galaxy only three supernova outbursts have been observed during the last thousand years, and in other galaxies they occur on the average at the rate of two per millennium. It is hardly possible that a refinement of the data will remove the contradiction which we have obtained.

The possibility that the planetary nebulae are formed in eruptions of ordinary novae is refuted by the fact that the masses of the envelopes ejected by novae are very small (not more than $10^{-4}\ M_{\odot}$), and their velocities of expansion are comparatively large (of the order of several hundred kilometres per second). However, it must be noted that nova outbursts take place fairly frequently (approximately 30 per year in the Galaxy). Hence it would be sufficient to maintain the existing number of planetary nebulae if only one eruption in three hundred resulted in the formation of a planetary nebula. It is possible, for example, that the planetary nebulae are formed in the eruptions of particularly "slow" novae or in repeated eruptions in regions of the Galaxy where there is an increased density of interstellar gas, which retards the ejected envelopes. However, it is best to say at present that the problem of the origin of the planetary nebulae is still far from being solved.

PART V.

NOVAE

Chapter 26. Nova outbursts and their interpretation

1. **Observational data.** As is well known, novae is the name given to stars which suddenly increase in luminosity by a factor of thousands or tens of thousands, afterwards slowly dying down. The flare-up of a nova usually takes place in the course of a few days, and it dies down over a period of several years. Finally, the star returns to a luminosity which differs little from the one it had before the outburst. At the time of maximum luminosity, the absolute magnitudes of novae are on the average -6^m. In its ordinary state, i. e. before the outburst and many years after it, a nova has an absolute magnitude of about $+5^m$ (with a fairly large dispersion). Thus the mean amplitude of variation of the luminosity of novae is 11^m. Some observational data for a number of novae are given in Table 20.

Table 20

Star		Amplitude of luminosity change	Absolute magnitude at maximum	Distance in parsecs	Velocity in km/sec
Nova Aurigae	1891	9	—5·3	800	
Nova Persei	1901	13	—8·4	480	700
Nova Aquilae	1918	12	—9·3	430	1300
Nova Cygni	1920	13·5	—8·9	1470	400
Nova Pictoris	1925	12	—7·3	500	70
Nova Herculis	1934	13	—5·5	230	170
Nova Lacertae	1936	13	—8·6	1350	1300
Nova Puppis	1942	16·5	—8·0	500	1100

Besides the change in luminosity of novae, very great changes take place in their spectra. The spectroscopic history of a nova is briefly as follows.

No observations have been made of the spectrum of any nova before its eruption, since the stars which become novae are very faint and not at all noticeable. An exception is formed by the spectrum of Nova

Aquilae 1918, which was found on a plate taken with an objective prism. However, it has not been possible to draw any definite conclusions concerning this spectrum. For the time from the outburst to the time of maximum luminosity, the spectra of novae are imperfectly known because of the rapidity of the flare-up. Only in a few cases has it been possible to obtain spectrograms for a few hours before the time of maximum, and, for some "slow" novae, for a few days before. These spectrograms show that, on the rising part of the luminosity curve shortly before maximum, novae have spectra which are usually of class A or F. The characteristic feature of these spectra is *the displacement of all the lines towards the violet* by an amount corresponding to a velocity of the order of some hundreds of kilometres per second. The velocities equivalent to the displacements of the lines for individual novae are given in the last column of Table 20.

Immediately after maximum luminosity is reached, broad bright bands suddenly appear on the red side of the absorption lines, lying roughly symmetrically with respect to the central frequencies. Originally the bright-line spectrum corresponds to class A, and then changes to class B. Simultaneously, the structure of the bright bands changes, and new absorption lines appear at their violet edges. Subsequently, as the star's luminosity decreases, the continuous spectrum and the absorption lines become fainter, and the bright-line spectrum changes from class B to class O.

Some months after the outburst, bright forbidden lines characteristic of the spectra of gaseous nebulae (including the N_1 and N_2 lines of "nebulium") appear in the spectrum of a nova. With the appearance of these lines, the nova enters the "nebular stage" of its development. The subsequent disappearance of the nebular lines coincides with the return of the star to its original luminosity. The spectrum of the star at this point belongs to the Wolf-Rayet type. It is usually supposed that the Wolf-Rayet stage is the last in the history of a nova. However, observations made on a number of novae during some decades after their outbursts have shown that the Wolf-Rayet spectrum is subsequently replaced by a spectrum of class O (sometimes with traces of emission).

During the first few years after the outburst, nebulae are visible around novae, similar in external appearance to the planetary nebulae. These nebulae expand with very high velocities and are then dispersed into space. The existence of such nebulae leaves no doubt that, in the eruption of a nova, the outer layers of the star are detached from it. We shall show below that the changes in the luminosity and spectrum of a nova are explained by the gradual movement of the detached envelope away from the star.

The stars which most resemble the typical novae are the **recurring novae.** Unlike ordinary novae, which, during the entire time they have been observed, have erupted only once, recurring novae erupt several times. A list of the recurring novae at present known is given in Table 21.

Table 21

Star	Years of outbursts	Limits of variation of stellar magnitude	
N Orionis	1677, 1750, 1892	6^m to	$> 11^m$
T Pyxidis	1890, 1902, 1920, 1941	6	14
U Scorpii	1863, 1906, 1935	9	>17
RS Ophiuchi	1898, 1933	4	12
T Coronae Borealis	1866, 1946	2	11
N Sagittae	1913, 1946	7	15
N Sagittarii	1901 ?, 1919	<7	14

The outbursts of recurring novae are quite similar to those of typical novae, but are on a smaller scale. This leads us to suppose that the typical novae also erupt many times, but that the intervals of time between the eruptions considerably exceed the period during which observations have been made. This supposition is confirmed by the statistical results of the Moscow astronomers B. V. KUKARKIN and P. P. PARENAGO. Having compared the time intervals between eruptions with the amplitudes of luminosity variation for recurring novae and nova-like variables, these authors came to the conclusion that, the greater the former quantity, the greater the latter, on the average. Extrapolating to typical novae the dependence obtained, B. V. KUKARKIN and P. P. PARENAGO found that, for an amplitude of luminosity variation of 11^m, the time interval between eruptions should be about 3000 years.

The spectra of recurring novae in the intervals between outbursts are of great interest, as being the spectra of these stars in their normal states (since the duration of the eruptions is considerably less than the interval between them). Unfortunately, the observational data on these spectra are few in number. Nevertheless, it can be asserted that, in the intervals between outbursts, the recurring novae are very hot stars, sometimes with emission features in their spectra. It is important to notice that the spectra of recurring novae before and after eruption are not essentially different. Since the typical novae are not fundamentally different from the recurring novae, this result can be extended to them. In other words, the spectra of the typical novae before eruption should be the same as they are many years afterwards, i. e. should belong to class O. This conclusion should be particularly emphasised, since it was thought until recently that cooler stars of class A are liable to become novae.

In turn, the **nova-like variables** are similar to the recurring novae; these are stars which resemble novae in their changes of luminosity and spectrum. The nova-like variables differ from the recurring novae not only in the smaller scale on which they exhibit the phenomena which characterise nova eruptions, but also by their less regular repetition. Among the nova-like variables are groups of stars of the types U Geminorum, Z Andromedae, etc.

Much more violent catastrophes than the eruptions of typical novae take place in the eruptions of supernovae. The absolute magnitudes of supernovae at maximum luminosity reach -15^m, and the amplitude of the change of luminosity apparently exceeds 20^m. The outbursts of supernovae are very rare phenomena. In the last thousand years, only three supernovae have erupted in our Galaxy: in 1054 in the constellation of Taurus, in 1572 in Cassiopeia, and in 1604 in Ophiuchus. The Crab Nebula, which is the outcome of the eruption, is now observed in the position of the supernova of 1054. The very great luminosities of the supernovae at maximum, which are comparable with the luminosity of an entire galaxy, enable us to discover supernovae in other galaxies without particular difficulty. The study of these supernovae has shown that their spectra consist of emission bands superposed on a continuous spectrum, the width of the bands corresponding to radial velocities of the order of a thousand kilometres per second. However, the identification of the bands is very uncertain, because of the great complexity of the spectrum.

Thus all the "erupting" stars considered above can be divided into four groups — the supernovae, the ordinary novae, the recurring novae, and the nova-like variables. The scale of the eruptions of the stars in each group is less than in the preceding one. In what follows, we shall be concerned mainly with the ordinary novae.

2. The explanation of the observations. The attention of astronomers has long been attracted by nova outbursts, and many hypotheses have been advanced to explain the observational data. However, the correct solution of the problem — the ejection of an envelope of the star — was found only a quarter of a century ago. Here we shall show that the movement of the detached envelope away from the star should lead to just such changes in luminosity and spectrum as are actually observed in novae.

At the moment of eruption, let an envelope be detached from the star whose optical thickness in the continuous spectrum is much greater than unity. As the envelope expands, its optical thickness will diminish, but, until it becomes of the order of unity, the envelope will serve as both a reversing layer and a photosphere. In this case, the surface temperature being approximately constant, the expansion of the envelope

results in an increase in the star's luminosity. In consequence of the approach of the part of the envelope that is towards the observer, the absorption lines will be displaced towards the violet end of the spectrum. Such a spectrum is in fact observed on the rising part of the luminosity curve of a nova.

At the moment when the maximum luminosity is reached, the optical thickness of the envelope in the continuous spectrum becomes of the order of unity. At this time the radiation directly from the star begins to reach the outer layers of the envelope, and bright lines appear in it. The reason for the appearance of these lines is the same as in the case of gaseous nebulae, i. e. *fluorescence*. The radiation in the lines reaches the observer not only from the part of the envelope which is approaching him, but also from that which is receding. It is not absorbed in the envelope, because of the Doppler effect. The width of the bright lines therefore corresponds to twice the velocity of expansion of the envelope. On the violet side of a bright line there is an absorption line which arises in the part of the envelope which is approaching the observer and screens off the star. A diagram of the origin of the spectral lines in the expanding envelopes of novae is given in Fig. 62.

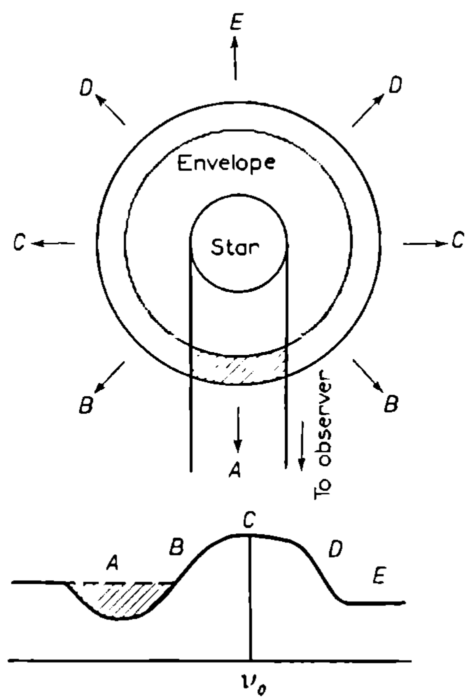

FIG. 62

After maximum luminosity, as the envelope expands further, its optical thickness in the spectral lines diminishes. As a result of this, the dark components of the bright lines become faint and then disappear. At the same time, the degree of excitation and ionisation of the atoms in the envelope increases, and owing to this the spectrum, as judged by its content of observed lines, passes from class A to class B, and then to class O. At some stage, the conditions necessary for the appearance of forbidden lines begin to be fulfilled in the envelope, i. e. the densities of matter and of radiation become fairly small. Beginning with the appearance of the forbidden lines, the "nebular stage" lasts for a fairly long time, until the brightness of the scattering envelope becomes less than that of the star itself. The spectrum of the star, which at this point is of Wolf-Rayet type, shows that the ejection of matter from the star is still continuing. When this process also comes to an end, the spectrum of the star becomes class O without emission lines.

It is clear that the continuous ejection of matter from the star begins immediately after the detachment of the envelope. In some cases further envelopes are detached from the star, and this leads to the appearance of secondary maxima on the falling part of the luminosity curve of the nova, and also to the appearance of additional systems of absorption lines. The process of ejection of matter from the star, which begins after the detachment of the main envelope, explains also a number of other features of the spectra of novae.

The explanations given above of the phenomena which occur when a nova erupts are, in their general features, in good agreement with observation for every stage of development of the nova. However, É. R. MUSTEL', who has studied in detail the observational data for the period around maximum luminosity, has recently come to the conclusion that two viewpoints (which he calls hypotheses A and B) are possible on the subject of the detachment of the envelope from the star.

According to hypothesis A, which has already been explained above, the detachment of the envelope from the star takes place at the time of the eruption. At first, the optical thickness of the envelope in the continuous spectrum is much greater than unity; later, it decreases and becomes of the order of unity at the time of maximum luminosity (the existence of the maximum is really due to this).

According to hypothesis B, an expansion of the entire star begins at the time of the eruption. At the time of maximum luminosity the envelope separates from the star, and the star itself begins to contract, in consequence of which its luminosity begins to decrease. The optical thickness of the detached envelope in the continuous spectrum is less than unity from the very beginning of the process. In other words, the detachment of the envelope takes place in the reversing layer of the star. É. R. MUSTEL' [96] gives many arguments against hypothesis A and in favour of hypothesis B. Unfortunately it is not possible, within the scope of a textbook, to pause to discuss this important work by É. R. MUSTEL'. It need only be remarked that at times long before or after the time of maximum luminosity there is almost no difference between the hypotheses A and B.

The explanation given above of the eruptions of novae allows us to devise simple methods of determining their parallaxes. The importance of these methods is the greater since the trigonometric parallaxes of novae are quite unreliable, because of their smallness.

One of the methods of determining the parallax of a nova, though not a very accurate one, is based on a comparison of the displacements of the absorption lines with the rate of increase of the luminosity before maximum. From observation we can find, for two instants t_1 and t_2,

the apparent magnitudes m_1 and m_2 and the temperatures T_1 and T_2 (from the spectral class). Using the well-known formula relating the absolute magnitude M of a star to its temperature T and radius R:

$$M = 29{,}500/T - 5 \log_{10} R - 0{\cdot}08 \,, \tag{26.1}$$

and the fact that the difference of the apparent magnitudes of the star is equal to the difference of its absolute magnitudes, i. e. $m_2 - m_1 = M_2 - M_1$, we obtain the following formula which determines the ratio of the star's radii at the instants t_1 and t_2:

$$\log_{10} \frac{R_2}{R_1} = \frac{5900}{T_2} - \frac{5900}{T_1} - \frac{m_2 - m_1}{5} \,. \tag{26.2}$$

On the other hand, we have for the difference of the radii of the star at the moments t_1 and t_2

$$R_2 - R_1 = v(t_2 - t_1) \,, \tag{26.3}$$

where v is the rate of expansion of the photosphere, found from the displacement of the absorption lines. Each of the quantities R_1 and R_2 is determined separately from (26.2) and (26.3). This makes it possible to find the absolute magnitude of the nova from the relation (26.1), and then to find the parallax by comparison with the apparent magnitude.

Another method of determining the parallax of a nova is based on the measurement of the rate of expansion of its envelope. This rate can be measured, on the one hand, from the width of the bright bands in the spectrum and expressed in kilometres per second, and on the other hand, from the observed expansion of the nebular envelope and expressed in angular measure. A comparison of these quantities gives the parallax of the nova. This method is more exact than the previous one. The distances and absolute magnitudes at maximum, given in Table 20 for a number of novae, were determined by this method.

The parallax of Nova Persei 1901 was found in an interesting manner. The nebula observed around this nova was expanding so fast that it could not be regarded as an envelope ejected in the eruption. This suggested that Nova Persei erupted inside a dust cloud and produced around itself an illuminated region which expanded with the velocity of light. This supposition was confirmed by the fact that the spectrum of the nebula obtained eighteen months after the outburst was the same as the spectrum of the star at the time of maximum light.

The parallax of Nova Persei 1901 was determined by the second of the methods mentioned above, taking into account that the velocity of "expansion" of the illuminated region was equal to that of light,

i. e. 300,000 km/sec. Later, a second nebula was discovered around Nova Persei, expanding much more slowly than the first. This was the "true" envelope which was detached from the star in the eruption.

3. **The interpretation of the luminosity curve.** Let us now consider the theoretical interpretation of the observational data. We shall first calculate the luminosity curve of a nova, assuming that, at the time of the eruption, there separates from the star an envelope of very great optical thickness in the continuous spectrum (i. e. assuming hypothesis A). Such a calculation was first performed by V. A. AMBARTSUMYAN, and in greater detail by SH. G. GORDELADZE [48].

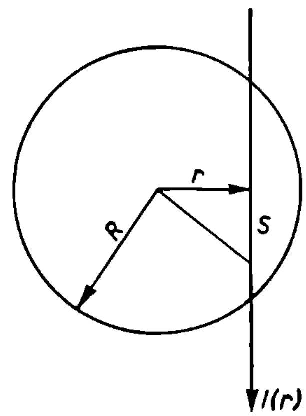

FIG. 63

The structure of the envelope of a nova is very complex, but as a first approximation we replace it by a homogeneous sphere with the same temperature throughout. The luminosity of such a sphere is very easily calculated. Since, according to our assumptions, ϱ = constant and T = constant, we can suppose that the emission coefficient ε and the absorption coefficient k are also constants. In this case, the intensity of radiation emerging at a distance r from the centre of the disc (Fig. 63) is

$$I(r) = \int\limits_{-\sqrt{(R^2-r^2)}}^{+\sqrt{(R^2-r^2)}} \varepsilon\, e^{-k[\sqrt{(R^2-r^2)}-s]}\, \mathrm{d}s\,,$$

or, effecting the integration,

$$I(r) = \frac{\varepsilon}{k}\,[1-e^{-2k\sqrt{(R^2-r^2)}}]\,. \qquad (26.4)$$

The total amount of energy emitted by the sphere, i. e. its luminosity L, is obtained from (26.4) by integrating over the whole disc and multiplying by $4\,\pi$:

$$L = 4\,\pi\,2\,\pi \int\limits_0^R I(r)\, r\, \mathrm{d}r\,.$$

The integration gives

$$L = 4\,\pi^2\,R^2\,\frac{\varepsilon}{k}\left[1+\frac{1}{k\,R}\,e^{-2kR}-\frac{1}{2\,k^2\,R^2}\,(1-e^{-2kR})\right].$$

Assuming that there is local thermodynamic equilibrium, we can put $\varepsilon/k = B$, where B is the intensity of radiation from a black body

at the given temperature. Further, we note that $k\,R$ is just the optical radius of the sphere. We denote it by τ_0. With this notation, we obtain

$$L = 4\,\pi^2\,R^2\,B\left[1 + \frac{1}{\tau_0}\,e^{-2\tau_0} - \frac{1}{2\,\tau_0^2}\,(1-e^{-2\tau_0})\right]. \qquad (26.5)$$

Formula (26.5) determines the luminosity of the sphere as a function of the radius R and the optical radius τ_0. However, it is easy to establish the relation between R and τ_0. To do so, we must write down the expression for the volume absorption coefficient k. We shall take it in the form

$$k = \beta\,\varrho^2/T^{9/2}\,,$$

where β is a constant. For τ_0 we have

$$\tau_0 = k\,R = \beta\,\varrho^2\,R/T^{9/2}\,,$$

or, expressing the density ϱ in terms of the mass M of the sphere and its radius R,

$$\tau_0 = 9\,\beta\,\mathsf{M}^2/16\,\pi^2\,T^{9/2}\,R^5\,. \qquad (26.6)$$

Substituting R from (26.6) in (26.5), we find

$$L = 4\,\pi^2\,B\left(\frac{9\,\beta}{16\,\pi^2}\right)^{2/5}\frac{\mathsf{M}^{4/5}}{T^{9/5}}\,f(\tau_0)\,, \qquad (26.7)$$

where

$$f(\tau_0) = \tau_0^{-2/5}\left[1 + \frac{1}{\tau_0}\,e^{-2\tau_0} - \frac{1}{2\,\tau_0^2}\,(1-e^{-2\tau_0})\right]. \qquad (26.8)$$

Let us now apply formula (26.7), which determines the luminosity of the sphere, to the envelopes of novae. Here we shall suppose that, when the envelope expands, its mass M and temperature T remain constant. Hence the luminosity L will depend only on the single variable τ_0. At the start of the eruption, $\tau_0 \gg 1$. Later, R increases, and consequently τ_0 decreases, by (26.6). For some value of τ_0, the luminosity L has a maximum. It is easily found that the maximum of the function $f(\tau_0)$ is reached for $\tau_0 = 1{\cdot}7$, and the maximum value of this function is

$$f_{\max} = 0{\cdot}84\,.$$

Hence we find for the luminosity of the nova at maximum

$$L_{\max} = 4\,\pi^2\,B\left(\frac{9\,\beta}{16\,\pi^2}\right)^{2/5}\frac{0{\cdot}84}{T^{9/5}}\,\mathsf{M}^{4/5}\,. \qquad (26\,9)$$

We see that, for a given temperature, the luminosity of a nova at maximum is the greater, the greater the mass of the ejected envelope. We shall use this important result below in determining the masses of the envelopes.

We can now rewrite formula (26.7) in the form

$$L = 1{\cdot}2\, L_{\max} f(\tau_0)\,. \tag{26.10}$$

The luminosity curve of the nova is determined by this formula, if we take into account the fact that the dependence of τ_0 on R is given by the relation (26.6), and that of R on the time t by the relation $R = v\,t$, where v is the rate of expansion of the outer boundary of the envelope. We see from formulae (26.5) and (26.6) that, when $\tau_0 \gg 1$, the luminosity increases as R^2, and when $\tau_0 \ll 1$ it decreases as R^{-3}. This means that the increase in luminosity occurs very rapidly, and the decrease considerably more slowly. Consequently the theoretical luminosity curve is similar in form to the observed one.

However, the above theory has some important defects. It cannot answer the following questions: (1) What is the source of energy whereby the ejected envelope radiates? (2) How does the transfer of energy through the envelope take place? These questions have been discussed by V. V. SOBOLEV [159], who considers two possible sources for the energy of radiation of the envelope: the energy in the envelope at the time when the eruption begins, and the energy emitted by the star after the envelope has broken away from it. The problem of the radiation of an expanding envelope with given sources of energy is then solved. This gives the variation in the luminosity and spectrum of a nova with time, in the interval between the beginning of the outburst and the time of maximum brightness. It was found that the theoretical results are in general agreement with those of observation. In particular, the theory accounts for the delay in the increase in brightness that has been observed in several novae.

4. The interpretation of the spectrum. We have already given a qualitative explanation of the spectra of novae, based on the idea of the expansion of an envelope ejected from the star. However, the quantitative interpretation of the spectra meets with great difficulties, in consequence of their extreme complexity. The spectra of novae in the nebular stage form an exception. In this case we can use the results obtained for the planetary nebulae. We shall therefore discuss here the spectra of novae in the nebular stage.

The nebular stage of a nova begins when the envelope has moved to such a great distance from the star that the conditions necessary for the appearance of forbidden lines begin to be fulfilled. It can be

shown that the condition concerning the density of radiation begins to be fulfilled sooner than that concerning the density of matter. Consequently, the beginning of the nebular stage is at the time when the density of free electrons in the envelope diminishes to such an extent that the spontaneous transitions in the forbidden lines become as frequent as collisions of the second kind. This enables us to estimate the density of free electrons in the envelopes of novae at the time when forbidden lines appear in their spectra, if the corresponding transition probabilities are known. On the other hand, we can estimate the volume of the envelope at the same epoch from the velocity of expansion of the envelope and the time interval which has elapsed since the eruption. This gives an approximate determination of the mass of the ejected envelope. It is found to be of the order of 10^{27} to 10^{28} grams. We shall later indicate more exact methods of determining the masses of the envelopes.

In the nebular stage of the nova, the electron temperature of the envelope can also be determined. This is most simply done by means of V. A. AMBARTSUMYAN's method, i. e. from the ratio of intensities of the forbidden lines $N_1 + N_2$ and 4363 Å, belonging to the O III ion. As is seen from formula (24.13), which determines the ratio of intensities of these lines, the former must always be brighter than the latter. We know that this is in fact the case in planetary nebulae. A similar ratio of intensities of the lines $N_1 + N_2$ and 4363 Å is observed in the spectra of novae also, when the conditions in the envelopes become similar to those in the nebulae. The electron temperatures of the envelopes at this time are found to be of the order of 6000° to 10,000°, like the temperatures of the nebulae.

However, at the time when the lines $N_1 + N_2$ and 4363 Å first appear in the spectra of novae, the ratio of their intensities is the reverse of what is observed in the spectra of the nebulae, i. e. the 4363 Å line is brighter than the $N_1 + N_2$ lines, which contradicts formula (24.13). The explanation of this phenomenon is as follows. In the derivation of formula (24.13) it was assumed that the number of spontaneous transitions in the forbidden lines is much greater than the number of collisions of the second kind (which is quite correct for the nebulae and for the envelopes of novae in their later stages). At the time of the first appearance of the forbidden lines in the spectra of novae, however, the number of collisions of the second kind is still comparable with the number of spontaneous transitions. Hence formula (24.13) is inapplicable at this time, and must be replaced by one which takes account of both spontaneous transitions and collisions of the second kind. Let us consider, for simplicity, the case where the collisions of the second kind predominate over the spontaneous transitions. In this case the transitions

of O III ions from the lower to the upper states in collisions of the first kind are almost completely balanced by transitions from the upper to the lower states in collisions of the second kind. The distribution of O III ions among states is therefore close to the Boltzmann distribution for the electron temperature of the gas in the envelope. Consequently, we have for the ratio of the numbers of O III ions in the $^1S_0$ and $^1D_2$ states

$$\frac{n_3}{n_2} = \frac{g_3}{g_2} e^{-h\nu_{23}/kT_e} .$$

This gives for the ratio of intensities of the lines $N_1 + N_2$ and 4363 Å

$$\frac{E_{N_1+N_2}}{E_{4363}} = \frac{A_{21}\,\nu_{12}\,g_2}{A_{32}\,\nu_{23}\,g_3} e^{h\nu_{12}/kT_e} . \tag{26.11}$$

For the O III ion, the coefficient A_{32} is more than 100 times A_{21}, and the ratio g_2/g_3 is 5. In the case considered, the ratio $E_{N_1+N_2}/E_{4363}$ is therefore greater than unity if $T_e < 10{,}000°$ and less than unity if $T_e > 10{,}000°$. Observation shows, as we know, that, at the time of appearance of the forbidden lines of the O III ion, the line 4363 Å is brighter than the $N_1 + N_2$ lines. Consequently the electron temperature of the envelope is comparatively high at this time.

Since the radiation of the envelopes of novae in the nebular stage is quite similar to that of the planetary nebulae, it is possible to apply Zanstra's method of determining the temperatures of the stars to novae also. This was first done for Nova Aquilae 1918. From the lines of He II it was found that three months after the eruption the temperature of the star was 65,000°. Subsequently Zanstra's method was applied to novae in their earlier stages. However, where the envelope is close to the star, the assumptions underlying Zanstra's method are not well satisfied (see Part VI).

The results obtainable by studying the contours of emission bands in the spectra of novae are very valuable. Since the envelopes are ejected with very high velocities, the contour of an emission band is determined mainly by the Doppler effect. For this reason, the distribution of velocities in the envelope can be estimated from the form of the contour. However, in the first period after maximum luminosity, when the envelope cannot yet be supposed transparent to the radiation in the lines, the theory of the contours of emission bands is very complex. It is considerably simpler in the subsequent period, when the absorption satellites of the emission bands disappear. The most reliable analysis is that for the forbidden lines, where self-reversal plays no part, since the absorption coefficient in forbidden lines is very small.

Let us assume that the envelope is transparent to radiation in a given line. If the envelope has spherical symmetry and all the layers are moving with the same velocity, the contour of the emission band will be flat, i. e. the intensity will be constant within the band. In some novae, the contours of the forbidden lines are in fact of this nature, at least to a first approximation. However, in other cases the bands seem to be doubled. This doubling was particularly marked in the spectrum of Nova Herculis 1934. The dispersion of velocities in a spherically symmetrical envelope cannot explain the doubling of the emission bands. We must therefore assume that the envelopes of some novae do not have spherical symmetry, i. e. the ejection of matter does not take place uniformly in different directions.

In connection with this result, it must be pointed out that individual condensations of matter are observed in the expanding nebulae around novae. Such an observation was first made in the case of Nova Pictoris 1925. At first it was interpreted as a division of the star into three components on eruption. These components moved in straight lines away from a single point, gradually diminishing in brightness. Six months after the outburst of Nova Herculis 1934, this star was discovered to be double. Later it was possible to obtain the spectra of each of the components, which were found to be characteristic of gaseous nebulae. This fact allowed the conclusion to be drawn that in reality, when a nova erupts, there takes place not a division of a single star into two or three components, but the ejection by the star of several condensations of matter, which radiate under the action of the star's radiation. In the case of Nova Herculis, this conclusion is well confirmed by the spectral data, since the doubling of the emission bands can be explained by the ejection of two condensations of matter from the star, which move with different radial velocities.

As has been said above, we shall not concern ourselves here with a detailed consideration of the spectra of novae during the period from maximum luminosity until the nebular stage. These spectra are very complex, and even a qualitative interpretation of them meets with difficulties. We notice, however, that at present a great part in the formation of the spectrum is ascribed not only to the main envelope, but also to the matter which is ejected from the star after the separation of the envelope, and that not only fluorescence but also collisions between the separate parts of the ejected matter are regarded as mechanisms for the appearance of the bright lines. Very valuable results from the study of the spectra of novae during this period have been obtained by É. R. Mustel' [98, 99, 101].

5. **Nova Herculis 1934.** The picture sketched above of the variation of the luminosity and spectrum of a nova is typical of the majority

of novae, but not all. Some novae show very considerable deviations from this picture. One of these "peculiar" novae is Nova Herculis 1934. The luminosity curve of this star is shown in Fig. 64. The luminosity, having first risen from 14^m or 15^m to $1^m.3$ (22 December), later decreased

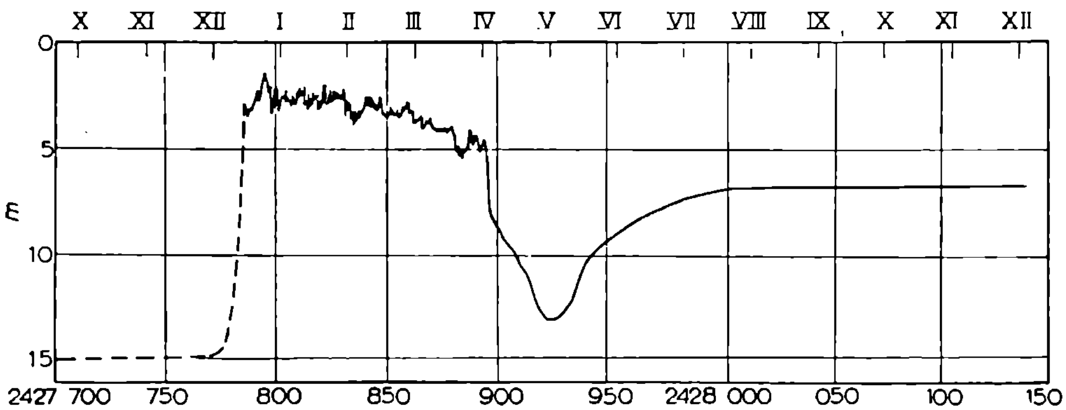

FIG. 64

slowly during a period of over three months. In April 1935 the luminosity of the nova suddenly fell rapidly to $13^m.1$, and then rose to approximately 7^m, after which it began to decrease slowly again. In the period of slow dying-down of the nova after the December maximum, its spectrum belonged to class F with emission lines of H, Fe II, Ca II, etc. The absorption lines were displaced to the violet by an amount corresponding to about 400 km/sec. After the April minimum, the spectrum of the nova became that typical of gaseous nebulae. The increase of the line intensities in this spectrum brought about the increase in brightness of the nova to 7^m.

The explanation of the changes in the luminosity and spectrum of Nova Herculis is as follows [52]. Starting from the moment of eruption, a powerful ejection of matter from the star took place for over three months, which led to the formation of an extended envelope around the star. However, the outermost parts of the envelope were not luminous during this period, since the ultra-violet radiation of the star did not reach them. This radiation was absorbed by the parts of the envelope which were closer to the star and denser, and these re-emitted it in the continuous spectrum with an energy distribution corresponding to a lower temperature than that of the star itself. In other words, the parts of the envelope adjoining the star played, during this period, the part of an extended photosphere. In April the intensity of outflow of matter diminished very greatly, the extended photosphere was dispersed, and a very hot star (with a temperature of about 70,000°) was revealed. Owing to the dissipation of the extended photosphere, the luminosity of the nova in the visible part of the spectrum decreased markedly.

Immediately afterwards, however, the outermost rarefied parts of the envelope began to radiate under the action of the ultra-violet radiation of the star, similarly to the radiation of the planetary nebulae, and for this reason the luminosity of the nova in the visible region of the spectrum began to increase again. This process of increase in luminosity continued until the intensity of the emission lines in the spectrum of the envelope corresponded to that of the ultra-violet radiation of the star. The further slow decrease of the luminosity of the nova was due to the gradual dissipation of the nebular envelope.

It is very interesting to consider the radiation of the envelope of Nova Herculis after the April luminosity minimum. In theoretical astrophysics, the assumption is usually made that radiative equilibrium exists in stellar envelopes. Even in those cases where the change in the physical conditions in the envelopes takes place very rapidly, it is still supposed that radiative equilibrium is established. In other words, the development of the envelope is imagined as a passage through a sequence of equilibrium states. However, we can now point to a series of examples of envelopes radiating in the absence of radiative equilibrium. One such example is the envelope of Nova Herculis after the April luminosity minimum. We have seen that, although the ultra-violet quanta from the star reached the nebular envelope at the time of minimum luminosity, it did not as yet radiate. Radiative equilibrium clearly did not exist at this time. In the following period, the development of the nebula was in the direction of the establishment of radiative equilibrium. This process can be regarded as having been completed only at the time of the secondary luminosity maximum.

Let us consider the theoretical interpretation of the radiation of the envelope of Nova Herculis after the April luminosity minimum [158]. For simplicity, we shall discuss a hydrogen envelope of constant density. We assume that the envelope has spherical symmetry, and that its thickness is considerably less than its distance from the star ($r_2 - r_1 \ll r_1$).

Let $n_1(r,t)$ and $n_+(r,t)$ be respectively the number of neutral and ionised atoms in 1 cm^3 at a distance r from the star at time t. Let n be the total number of hydrogen atoms in 1 cm^3, so that

$$n_1(r,t) + n_+(r,t) = n \, . \tag{26.12}$$

At the initial instant, which we take as the instant when the temperature of the star suddenly increases, all the atoms in the envelope are neutral, i. e. $n_1(r,0) = n$, $n_+(r,0) = 0$.

Subsequently, ionisation of the atoms takes place by the absorption of quanta emitted by the star beyond the limit of the Lyman series. The number of such quanta incident on 1 cm^2 of the inner boundary

of the envelope in 1 second is denoted by H. If we denote by k the absorption coefficient, referred to one atom, the total number of ionisations occurring in 1 second in 1 cm³ at a distance r from the star at time t is

$$n_1(r,t)\, k\, H\, e^{-\tau(r,t)}\,,$$

where $\tau(r,t)$ is the optical distance of the point concerned from the inner boundary of the envelope, i. e.

$$\tau(r,t) = \int_{r_1}^{r} n_1(r,t)\, k\, \mathrm{d}r\,. \tag{26.13}$$

Besides ionisations, the converse processes, i. e. recombinations, also take place in the envelope. The number of recombinations to the ith level occurring in 1 second in 1 cm³ at a distance r from the star at time t is

$$n_e(r,t)\, n_+(r,t)\, C_{if}(T_e)\,,$$

where $n_e(r,t)$ is the number of free electrons in 1 cm³, and $C_{if}(T_e)$ is the function determined by formula (24.14).

Since the variation in the number of ions is equal to the difference between the numbers of ionisations and recombinations, we have

$$\partial n_+/\partial t = n_1\, k\, H\, e^{-\tau} - n_e\, n_+ \sum_{i=2}^{\infty} C_{if}\,. \tag{26.14}$$

In equation (26.14) we have not taken into account recombinations to the first level or ionisations under the action of the diffuse radiation of the envelope, since these processes balance each other.

Thus the problem of determining the variation in the number of ionised atoms in the envelope reduces to the solution of equation (26.14) with the conditions (26.12) and (26.13). The solution of this equation will be found in the work of V. V. Sobolev [158]. Here we shall give one result which will be needed in what follows.

Let the optical thickness of the envelope beyond the limit of the Lyman series, at the initial instant, be much greater than unity, i. e. $\tau(r_2,0) \gg 1$. Next, let the radiation of the star be so intense that it can produce in the nebula a degree of ionisation considerably greater than unity ($n_+/n_1 \gg 1$). In this case the envelope can be divided, at any time, into two regions: an "ionised" region (where $n_+/n_1 \gg 1$) and an "un-ionised" region (where $n_+/n_1 \ll 1$), with a very sharp boundary between them, and the process of the gradual increase in ionisation in the envelope can be regarded as a movement of the boundary between these regions. This boundary lies approximately where $\tau \approx 1$ at the

time considered (Fig. 65). This result is quite evident, since, so long as $\tau \gg 1$ for a given layer, the ionising radiation of the star does not penetrate to it. Only when, owing to the ionisation of the part of the envelope towards the star, its optical thickness becomes of the order of unity, does ionisation begin in the layer considered.

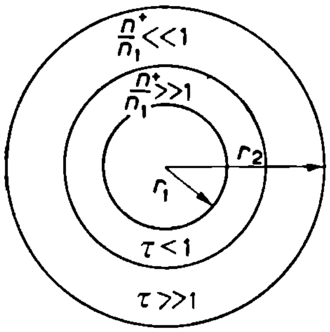

FIG. 65

This result enables us to find without difficulty how the total number of ionised atoms in the envelope, i. e.

$$N_+ = \int n_+ \, \mathrm{d}V ,$$

varies with time. We integrate both sides of equation (26.14) over the whole volume of the nebula. Denoting by $\mathfrak{N}$ the total number of quanta emitted by the star beyond the limit of the Lyman series in 1 second, we obtain

$$\mathrm{d}N_+/\mathrm{d}t = \mathfrak{N} - N_+ \, n \sum_{i=2}^{\infty} C_{if} . \tag{26.15}$$

The last term in equation (26.15) is obtained using the relation

$$\int n_e \, n_+ \, \mathrm{d}V = n \, N_+ . \tag{26.16}$$

This follows since for the ionised region $n_e \approx n$, and for the un-ionised region $n_e \approx 0$.

The solution of equation (26.15) which satisfies the initial condition $N_+(0) = 0$ has the form

$$N_+(t) = \mathfrak{N} \left[1 - \exp\left(-t n \sum_{i=2}^{\infty} C_{if}\right)\right] / n \sum_{i=2}^{\infty} C_{if} . \tag{26.17}$$

A very important result follows from formula (26.17). The time

$$t_* = 1/n \sum_{i=2}^{\infty} C_{if} \tag{26.18}$$

can be regarded as the time taken to establish radiative equilibrium (or the relaxation time of this process). We see that it is the greater, the smaller the density of matter in the envelope.

In order to estimate the relaxation time, we notice that for hydrogen $\sum_{i=2}^{\infty} C_{if} \approx 3 \times 10^{-13}$ (for $T_e = 10{,}000°$). This means that when $n \approx 10^{10}$ the relaxation time is of the order of several minutes. In the envelopes of novae in the first period after maximum luminosity, $n > 10^{10}$. We

can therefore suppose that the establishment of radiative equilibrium in this case follows immediately upon a change in the physical conditions in the envelope. However, in the case of Nova Herculis 1934, the temperature of the star suddenly increased when the density of matter in the envelope was very small. For this reason the time for the establishment of radiative equilibrium in the envelope of Nova Herculis was prolonged to approximately a month.

Using formula (26.17), we can find the variation in the total amount of energy emitted by the envelope in any spectral line. We have for the energy emitted by the envelope in a line of frequency ν_{ik}

$$E_{ki} = A_{ki}\, h\nu_{ik}\, z_k \int n_e\, n_+\, \mathrm{d}V\,, \tag{26.19}$$

where $z_k = n_k/n_e\, n_+$. Using formulae (26.16) and (26.17), we obtain

$$E_{ki} = A_{ki}\, h\nu_{ik}\, z_k\, \Pi\, [1 - \exp\,(- t\, n \sum_{i=2}^{\infty} C_{if})] / \sum_{i=2}^{\infty} C_{if}\,. \tag{26.20}$$

For Nova Herculis, formula (26.20) can be compared with observation. It is found that the observed variation in the intensities of the Balmer lines is fairly well represented by formula (26.20). Here it is necessary to take the value $n = 3 \times 10^6$ for the concentration of atoms in the envelope.

Using the formulae given above, we can also explain an interesting fact in the spectroscopic history of Nova Herculis. The red components of the emission lines were, at the time when they appeared after the April luminosity minimum, much fainter than the violet components, and the two were approximately comparable in intensity only at the time of the secondary luminosity maximum. To explain this phenomenon, it was previously assumed that, in the space between the star and the envelope, there was absorbing matter, which caused the red components of the emission lines to become fainter. Now, however, it can be more naturally explained in terms of the radiation of the envelope in the absence of radiative equilibrium.

It follows from the observations that the intensity of the red component increased more slowly than that of the violet component. On the above view regarding the radiation of Nova Herculis during this period, this means that radiative equilibrium was established more slowly in the receding part of the envelope than in the approaching part. For this to be so, we need only assume that the density of matter in the receding and approaching parts was not the same. The phenomenon is described approximately quantitatively by the above formula (26.20).

This conclusion that the envelope of Nova Herculis was inhomogeneous is confirmed by observational results. It is well known that there were visible in the envelope two well-defined concentrations of matter, one receding and the other approaching. This, indeed, would also explain the two-component structure of the emission lines in the spectrum of this nova. If we started from the presence of two concentrations of matter and supposed that in this period the envelope was not in radiative equilibrium, we should have to conclude that the intensities of the red and violet components of the emission lines would vary differently, since it is difficult to imagine that the two concentrations would be completely similar.

Chapter 27. The part played by the ejected envelopes in the evolution of novae

1. The masses of the envelopes. The subject of the masses of the envelopes ejected in the eruptions of novae is of great interest. Firstly, from the masses of the envelopes we can obtain information about the layers in which the detachment of the envelope from the star takes place, and about the forces which cause this detachment. Secondly, a knowledge of the masses of the envelopes allows us to draw conclusions about their cosmogonical significance. If the mass of the envelope is comparable with that of the star, the eruption must bring about a radical change in the entire star. On the other hand, if the mass of the envelope is small compared with that of the star, the change in the star occurring as a result of the eruption is evidently restricted to the surface layers alone.

Several methods have been proposed to determine the masses of the envelopes. We shall give three of these.

(1) *Determination of the mass of the envelope from the maximum luminosity.* This method is based on formula (26.9), which was obtained in interpreting the luminosity curve of a nova. Formula (26.9) asserts that, the greater the mass of the ejected envelope, the greater the maximum luminosity of the nova. Converting the luminosity of the star in formula (26.9) to the absolute magnitude, and putting $T = 10{,}000°$, we obtain

$$M_{\max} = -17{\cdot}2 - 2\log_{10}(\mathsf{M}/\mathsf{M}_\odot)\,, \tag{27.1}$$

where $\mathsf{M}_\odot$ is the mass of the Sun. Since the absolute magnitudes of novae at maximum are in the neighbourhood of -6^{m}, the masses of the envelopes, calculated by formula (27.1), are of the order of 10^{-6} to $10^{-4}\ \mathsf{M}_\odot$.

(2) *Determination of the mass of the envelope from the luminosity and volume in the nebular stage.* When the envelope of a nova is in the nebular stage, its mass can be determined by the same method as is used to determine the masses of planetary nebulae. The mass of the envelope is approximately

$$\mathsf{M} = C \sqrt{(L V)}\,, \tag{27.2}$$

where C is some constant, L is the luminosity, and V is the volume of the envelope. The determination of the masses of the envelopes of novae from formula (27.2) again leads to values of the order of 10^{-6} to 10^{-4} $\mathsf{M}_{\odot}$. However, for individual novae this method gives more exact results than the previous one.

(3) *Determination of the mass of the envelope from the relaxation time.* This method can be applied only to novae which are similar to Nova Herculis 1934. We have already seen that in this case the radiation of the envelope before the secondary luminosity maximum takes place without the establishment of radiative equilibrium. Having found, from the luminosity curve, the time for the establishment of radiative equilibrium (the relaxation time), and using formula (26.18), we can determine the density of the envelope. On the other hand, a knowledge of the rate and duration of the outflow of matter from the star enables us to estimate the volume of the envelope. The mass of the ejected envelope is thus obtained.

In order to avoid having to estimate the volume of the envelope, which involves some uncertainty, we can proceed as follows. We rewrite the relation (26.18) in the form

$$t_* \, n_e \sum_{i=2}^{\infty} C_{if} = 1$$

(we have replaced n by n_e, since $n_e = n_+ \approx n$). Multiplying both sides of this equation by n_+ and integrating over the whole volume of the envelope, we obtain

$$t_* \int n_e \, n_+ \sum_{i=2}^{\infty} C_{if} \, \mathrm{d}V = \int n_+ \, \mathrm{d}V \,.$$

The integral on the right-hand side of this equation, however, is the total number of hydrogen ions in the envelope, and that on the left-hand side is the total number of recombinations to all levels from the second upwards. The latter number is equal to the number of Balmer quanta emitted by the envelope in 1 second. Consequently we have

$$t_* \, N_{\mathrm{B}} = N_+ \,. \tag{27.3}$$

Formula (27.3) holds for any time before the secondary luminosity maximum. From the relaxation time t_* and the number N_B of Balmer quanta emitted by the envelope, it gives the number of hydrogen ions N_+ in the luminous (i. e. the ionised) region. Since, in the ionised region, $n_+/n_1 \gg 1$, N_+ is almost the same as N, i. e. the total number of hydrogen atoms in this region. In the course of time, the dimensions of the ionised region increase, and become a maximum when radiative equilibrium is established. The number N, determined for this time from formula (27.3), signifies either the total number of hydrogen atoms in the envelope (if the latter is completely ionised) or a lower limit to this number (if the envelope is only partly ionised). For Nova Herculis 1934, at the time of the secondary luminosity maximum, formula (27.3) gives $N = 1{\cdot}4 \times 10^{52}$, and therefore $\mathsf{M} = m_{\mathrm{H}} N = 2{\cdot}3 \times 10^{28}$ g.

Thus we see that all these methods lead to values for the masses of the envelopes of novae of the order of 10^{-6} to 10^{-4} times the Sun's mass. The different values which are obtained for individual novae by using the different methods are explained not only by the inexactness of these methods, but also by the fact that they are applied to different times. In the course of time, however, owing to the continuous ejection of matter from the star, the mass of the envelope increases. Hence the second and third of the methods given above should give higher values for the masses than the first.

However, it is necessary to remark that none of the above methods gives a complete idea of the total amount of matter ejected as a result of the eruption of a nova, since the outflow of matter from the star is continuing even when the nova has reached its minimum luminosity. We deduce this from the fact that at this time the nova has a spectrum of the Wolf-Rayet type. The time during which the nova remains in the Wolf-Rayet stage may be estimated as some decades. The amount of matter ejected by the star in this time is approximately equal, in order of magnitude, to that which it loses during the eruption itself.

Thus we find that the amount of matter ejected as a result of the eruption of a nova is a very small part of the mass of the star. This means that the eruption of a nova does not bring about any noteworthy changes in the structure of the star. At the same time, the smallness of the mass ejected in the eruption is in agreement with the hypothesis that some stars undergo eruption many times.

It is also very important to estimate the amount of matter ejected in the eruption of a supernova. The immense brightness of supernovae at maximum compels us to suppose that in this case a mass is ejected which much exceeds that ejected in the eruption of an ordinary nova. Formula (27.1) gives some idea of the masses of the envelopes of supernovae. Substituting in this formula the value $M_{\max} = -15^{\mathrm{m}}$ for the

absolute magnitude at maximum, we find that the mass of the envelope is $M = 0 \cdot 1\ M_{\odot}$. However, there is no certainty that formula (27.1) is applicable to supernovae.

We can more reliably estimate, from formula (27.2), the mass of the Crab Nebula, which, as is well known, was ejected in the eruption of the supernova of 1054. However, it must be borne in mind that the spectrum of the Crab Nebula is rather different from those of the planetary nebulae and of the envelopes of novae in the nebular stage. The emission lines of hydrogen are very faint in it, but the continuous background, on the other hand, is strong. It seems that in this case some peculiar mechanism of excitation is involved. For this reason, the value of the constant C in this case will be somewhat different from its value for the planetary nebulae. With this reservation, we find from formula (27.2) that the mass of the Crab Nebula is of the order of two solar masses.

The estimates given show that, in the eruption of a supernova, a mass is ejected which comprises a considerable part of the mass of the star.

2. **The dynamics of the envelopes.** Observation shows that the displacements of the absorption lines in the spectra of novae undergo considerable changes in some cases. This fact, interpreted as a change in the rate of motion of the envelope, enables us to draw conclusions concerning the forces which act on it.

However, it must be remarked at once that the acceleration or retardation of the envelopes is not the only possible explanation of the observed changes in the displacement of the absorption lines. Another similar effect is due to the variation in the effective level of the absorbing material in an envelope moving with a velocity gradient. Let us assume, for instance, that the outer layers of the envelope are expanding with a greater velocity than the inner ones. In this case, as the outer layers become dispersed, the effective level of the absorbing matter will approach the inner boundary of the envelope, and the displacement of the absorption lines will decrease. Yet another cause of this effect may be the influence of the matter ejected from the star after the luminosity maximum, since the rate of motion of this matter is, generally speaking, not the same as that of the main envelope. These facts compel us to proceed with caution in drawing conclusions regarding the dynamics of the envelopes from the observed displacements of absorption lines.

In the period from the time of eruption to the time of maximum luminosity, a decrease in the displacement of absorption lines has been observed in the spectra of some novae. A number of authors have explained this effect by the retardation of the envelopes under the action of gravity from the star; this enables us to determine the masses

of the novae. The most complete investigation of this topic has been given by É. R. MUSTEL' [101]. He has found that the masses of a number of novae lie between 70 $M_\odot$ (Nova Cygni 1920) and 1300 $M_\odot$ (Nova Aquilae 1918). Such large masses are not found in any of the classes of star known to us. Of course, it may be that such unusual stars as novae have unusually large masses. However, it is not impossible that the decrease in the displacement of the absorption lines in this case is in part due to the phenomenon, mentioned above, of a change in position of the effective level of the absorbing matter.

Concerning the question of the masses of novae, it must be remarked that the conclusion that the masses are very large is an inevitable consequence of "hypothesis B", proposed by É. R. MUSTEL' (see Chapter 26). As we know, according to this hypothesis an expansion of the entire star begins at the moment of the eruption, and the outer layers of the star are separated from it at the time of maximum luminosity; the photosphere remains with the star. Consequently, at the time of maximum luminosity the rate of expansion of the photosphere must be less than the parabolic velocity at a distance from the centre of the star equal to the radius of the photosphere at that time.

In other words, we must have

$$v_p < \sqrt{(2G\, M_*/r_p)}\,,$$

whence

$$M_* > v_p{}^2\, r_p/2\, G\,. \tag{27.4}$$

Having determined v_p and r_p by the same method as in finding the parallax of a nova, we can obtain from the inequality (27.4) a lower limit to the mass of the star. É. R. MUSTEL', using the inequality (27.4), found the same large values for the masses of the novae as from the decrease in the displacement of the absorption lines. For instance, the mass of Nova Aquilae 1918 was found to be 1700 $M_\odot$.

The most serious argument against large masses for the novae has been given by É. R. MUSTEL' himself. As is well known, novae have very small radii (of the order of one-tenth of the Sun's radius) at the time of minimum luminosity. We should therefore observe a red-shift in the spectra of novae, if their masses were large. However, observation does not reveal this. Thus the extremely important problem of the masses of the novae cannot yet be regarded as solved.

Immediately after the luminosity maximum, the displacements of the absorption lines in the spectra of novae greatly increase (for example, from 70 to 285 km/sec in the spectrum of Nova Pictoris 1925). At first, additional lines appear, displaced to the violet of those observed before the luminosity maximum. The intensity of the new lines gradually increases, and that of the old lines decreases, until they finally disappear

completely. This gives the impression that the second spectrum is formed at the expense of the first one. É. R. MUSTEL' explains this transformation of one spectrum into another by the action of selective radiation pressure. He thinks that, immediately after the luminosity maximum, atoms in the layers of the envelope nearest the star begin to absorb the photospheric radiation in the spectral lines, and this causes an increase in the velocity of these atoms. Owing to the Doppler effect, the layers of the envelope nearest the star become transparent to the photospheric radiation in these frequencies, which is then absorbed by the atoms in the next layers of the envelope. This process continues until the whole envelope acquires an increased velocity. Other hypotheses also have been proposed to explain the increase in the displacement of absorption lines in the spectra of novae after the luminosity maximum, but that of É. R. MUSTEL' is the most plausible.

It is certain that the two forces mentioned above (the attraction of the star and the force of radiation pressure) play a large part in the dynamics of the envelope when it is close to the star. At a later stage, however, the retardation of the envelope by the resistance of the interstellar medium may be of importance. Let us briefly consider this effect.

We assume that the nova has erupted in a homogeneous medium of density ϱ. As the envelope expands, particles of this medium will be incident on it, and its mass will increase. If the mass of the envelope at the time of the eruption was M, at a distance r from the star it will become

$$\tfrac{4}{3}\,\pi\, r^3\, \varrho + \mathsf{M}\,.$$

From the law of conservation of momentum, we can write

$$(\tfrac{4}{3}\,\pi r^3 \varrho + \mathsf{M})\, v = \mathsf{M}\, v_0\,, \tag{27.5}$$

where v_0 is the velocity of the envelope at the initial instant and v is its velocity at a distance r from the star. Substituting $\mathrm{d}r/\mathrm{d}t$ for v in equation (27.5) and integrating, we obtain

$$\tfrac{1}{3}\,\pi\, r^4\, \varrho + \mathsf{M}\, r = \mathsf{M}\, v_0\, t\,, \tag{27.6}$$

where t is the time since the eruption. The relation (27.6) determines the radius r of the envelope as a function of the time t.

In order to find how the rate of expansion of the envelope varies with time, we must use the relations (27.5) and (27.6). Let us find, for instance, the time during which the velocity decreases by a factor of two. We have from (27.5)

$$v/v_0 = \mathsf{M}/(\tfrac{4}{3}\,\pi r^3 \varrho + \mathsf{M})\,. \tag{27.7}$$

It is evident that v will be equal to $\frac{1}{2} v_0$ when

$$\frac{4}{3} \pi r^3 \varrho = \mathsf{M} . \tag{27.8}$$

Substituting (27.8) in (27.6), we obtain for the required time

$$t = \frac{5}{4 v_0} \left(\frac{3 \mathsf{M}}{4 \pi \varrho}\right)^{1/3} . \tag{27.9}$$

Table 22 gives the time intervals during which the velocity of the envelope diminishes by factors of 2 and of 100, and also the radii of the envelope when these velocities are reached. For the density of the interstellar medium we take its mean value $\varrho = 3 \times 10^{-24}$ g/cm^3, and for the initial velocity of the envelope $v_0 = 1000$ km/sec; t and r are calculated from formulae (27.7) to (27.9). The table has been compiled for two values of the mass of the envelope, namely 10^{-5} and 10^{-4} solar masses.

Table 22

	$\mathsf{M} = 10^{-5} \mathsf{M}_\odot$		$\mathsf{M} = 10^{-4} \mathsf{M}_\odot$	
	$v/v_0 = 1/2$	$v/v_0 = 1/100$	$v/v_0 = 1/2$	$v/v_0 = 1/100$
t in years	48	4500	102	9800
r in parsecs	0·04	0·18	0·08	0·38

We see that the retardation of the envelopes of novae should become noticeable after some decades. However, this is not generally observed. For example, the envelope of Nova Aquilae 1918 has been expanding for over 30 years without any retardation. The absence of any noticeable retardation in this case is apparently explained by the comparatively large mass of the envelope of this nova (equal to $10^{-4} \mathsf{M}_\odot$). Another possible explanation is that, during the time interval between eruptions, the nova does not succeed in leaving the region from which the interstellar matter has been removed by the previous eruption.

If the eruption of a nova took place at a point where the density of interstellar matter is increased, the discovery of a retardation of the envelope would become more probable. In this respect Nova Persei 1901 is very interesting. It erupted, as we know, inside a dust cloud and illuminated it. J. H. Oort has compared photographs of the envelope of this nova taken in 1917 and 1934, and has discovered that during this time the motion of the envelope became slower, and at some points it was deformed. He explains this latter fact by an inhomogeneity of the dust cloud. It is interesting that the deformed edge of the envelope is very bright. Oort thinks that its radiation is due to

collisions between the atoms in the envelope and the particles in the dust cloud. OORT sees in this an additional confirmation of the retardation of the envelope.

3. **The energy evolved in the eruption.** We have already determined the mass ejected in the eruption of a nova. Let us now calculate the energy evolved in the eruption. This energy is composed of three parts: (1) radiant energy, (2) the kinetic energy of the envelope, and (3) the energy of separation of the envelope from the star. We shall consider each of these parts separately.

The *radiant energy* is determined from the formula

$$E_r = \int L(t)\,\mathrm{d}t\,, \tag{27.10}$$

where $L(t)$ is the luminosity of the nova, and the integration is extended over the entire period of the eruption. The integral (27.10) can be calculated for each nova by means of the luminosity curve. It is found that

$$E_r \approx 10^{45} \text{ to } 10^{46} \text{ ergs.}$$

The *kinetic energy* of the envelope is

$$E_k = \tfrac{1}{2}\, \mathsf{M}\, v^2\,. \tag{27.11}$$

Assuming that the masses of the envelopes lie between 10^{28} and 10^{29} g, and that their velocities are of the order of 1000 km/sec, we find that $E_k \approx 10^{44}$ to 10^{45} ergs.

To calculate the *energy of separation of the envelope* from the star, we must use the formula

$$E_s = G\, \mathsf{M}_*\, \mathsf{M}/R\,. \tag{27.12}$$

For the masses of the novae, we take values of the order of several solar masses. The radius of a nova at the time when the envelope is separated will be taken equal to the radius after eruption, i.e. $R = 0{\cdot}1\, R_\odot$. In this case, formula (27.12) gives $E_s \approx 10^{44}$ to 10^{45} ergs.

Thus we find that the total energy evolved in the eruption of a nova is 10^{45} to 10^{46} ergs. For comparison, it may be mentioned that the Sun emits this amount of energy in 10^5 to 10^6 years.

It is very important to know in which layers the release of energy takes place in the eruption of a nova. Two hypotheses may be advanced: (1) the energy is evolved in the deepest layers of the star and transmitted to the exterior in some manner, (2) the energy is evolved in the surface layers of the star, lying directly beneath the ejected layers.

It is easy to show that, in the first case, the transfer of energy cannot take place by radiation. For, if this were the case, at least a million years would be required for the energy to traverse the whole thickness of the star. Furthermore, the whole process of the emergence of the energy to the exterior would also be extended to a time of the order of a million years. Consequently, an explosion taking place in the central regions of the star would be observed as a very small increase in the luminosity over an extremely long period of time. None of the phenomena which characterise the outbursts of novae would, of course, be observed here.

For this reason, another mechanism of transfer of explosive energy from stellar centres to the exterior has been closely considered, namely the propagation of a shock wave. L. É. Gurevich and A. I. Lebedinskiĭ [54] have shown that, in this case, the transfer of the energy of the explosion to the surface of the star takes only about half an hour. A shock wave which is propagated in a medium with decreasing density moves with an acceleration, and may emerge at the surface with a velocity of some thousands of kilometres per second, though its velocity was originally much smaller.

However, the second hypothesis mentioned above appears more probable, namely that the explosion takes place where the envelope is detached from the star. If we were well acquainted with the structure of the star's photosphere before the eruption, we could use the known mass of the envelope to determine the physical conditions at the place of its detachment. Unfortunately, our knowledge of the state of novae before eruption is still insufficient for such calculations.

4. **The part played by the eruption in the evolution of the star.** The hypothesis was advanced by E. A. Milne that the eruption of a nova indicates the transformation of an ordinary star into a white dwarf. The basis for this hypothesis was the fact that a nova after eruption is a dwarf star (with an absolute magnitude of about $+5^{m}$) of Wolf-Rayet spectral type. However, if the transition is actually made, in the eruption of a nova, from one equilibrium configuration (an ordinary star) to another (a white dwarf), a very large amount of energy would have to be released in the process, approximately equal to the difference of the gravitational energies of these configurations. In order of magnitude this energy is

$$\Delta E = G\, \mathsf{M}_*^2 \left(\frac{1}{R_2} - \frac{1}{R_1}\right), \tag{27.13}$$

where M_* is the mass of the star, and R_1 and R_2 the initial and final radii of the configuration. For a mass of the nova equal to several solar masses, formula (27.13) gives $\Delta E \approx 10^{50}$ ergs. However, we have

seen that, in the eruptions of ordinary novae, an amount of energy of the order of 10^{45} to 10^{46} ergs is liberated, i. e. 10^4 to 10^5 times less than would be required on MILNE's hypothesis. This contradiction could be removed by making the assumption that the gravitational energy released goes to increase the other forms of energy in the star (for example, to increase the temperature of its interior, etc.). However, such an exact balance of these energies, as a result of which only a negligibly small part of the energy liberated is actually evolved, appears very improbable.

Another objection to MILNE's hypothesis follows from the statistics of novae. It is known that, on the average, 30 novae erupt every year in the Andromeda Nebula. The number of novae erupting every year in our Galaxy is apparently still greater, but an exact estimation of this number is hindered by the fact that only those novae are observed which erupt in the neighbourhood of the Sun. Assuming that the mean age of the stars is 10^{10} years, we find that at least 3×10^{11} novae have erupted in the Galaxy during its lifetime. The total number of stars in the Galaxy, however, is approximately 3×10^{10}. Consequently, each star has erupted, on the average, at least 10 times. On the other hand, we know for certain that the Sun has not undergone such a catastrophe in 2×10^9 years, since this would have led to the fusion of the Earth's crust, which has not occurred during that period. It may be supposed that other stars similar to the Sun have also not undergone eruptions during a time of the same order. This increases the number of eruptions occurring in each of the remaining stars. These facts compel us to suppose that there is a special class of stars, each of which erupts a very great number of times. We recall that the same conclusion was drawn above, on the basis of the similarity between typical and recurring novae.

These considerations, which disprove MILNE's hypothesis, have led to its replacement by the hypothesis that an ordinary star is transformed into a white dwarf not as a result of one nova eruption, but by means of several such eruptions. This hypothesis was developed, in its most general form, by B. A. VORONTSOV-VEL'YAMINOV [173].

First of all, B. A. VORONTSOV-VEL'YAMINOV pointed out that different novae have very different luminosities in the intervals between eruptions (i. e. in their normal state); in the spectrum-luminosity diagram, they occupy the region between the Wolf-Rayet stars and the blue and white dwarfs. A more detailed consideration of the question led to the discovery of a new continuous sequence in the spectrum-luminosity diagram, called by B. A. VORONTSOV-VEL'YAMINOV the **white-blue sequence**. This sequence begins at the O and B-type stars, passes through the Wolf-Rayet type stars and the recurring and ordinary

novae, and ends with the blue and white dwarfs. Since the masses of stars of class O are about 40 $M_\odot$, those of Wolf-Rayet stars of the order of 10 $M_\odot$, and those of the white dwarfs of the order of $M_\odot$ or less, B. A. VORONTSOV-VEL'YAMINOV made the assumption that the white-blue sequence has an evolutionary significance; the most massive stars, in losing their mass, move during their evolution along the sequence, ultimately becoming white dwarfs. The loss of mass takes place initially in the form of a continuous outflow, and later as a result of explosions, the intervals between which gradually increase.

The very fact of the existence of the white-blue sequence in the spectrum-luminosity diagram is undoubtedly of great interest. However, the hypothesis of the evolutionary significance of this sequence cannot be accepted at present. The following objections have been raised against it [102]:

(1) Let us assume that stars of class O are transformed into Wolf-Rayet stars, then into novae. and finally into white dwarfs. Since the Wolf-Rayet stars and the novae lose mass very rapidly, their lifetime must be comparatively short (of the order of 10^5 to 10^6 years for Wolf-Rayet stars, and of the order of 10^7 to 10^8 years for novae). It is clear that in such a short time neither the distribution of the stars in space nor their velocity distribution can noticeably change. Consequently, all the objects which appear in the white-blue sequence should have approximately the same spatial and kinematic characteristics. Observation, however, contradicts this conclusion.

(2) Let us consider the companions of stars appearing in the white-blue sequence. It is well known that O and Wolf-Rayet stars are very often found to be double, their companions being hot giants. Novae either do not have companions at all, or the latter are very faint stars. The white dwarfs are companions of very varied stars (for example, one of them is the companion of Sirius, another the companion of the long-period variable Mira Ceti, and so on). It is clear that, if we assume that the stars appearing in the white-blue sequence evolve along it, it becomes very difficult to explain the evolution of their companions.

Thus we can at present say nothing definite concerning the part played by eruptions in the evolution of the star. We merely emphasise once more that the stars which are subject to eruptions are markedly different in nature from the other stars. For example, stars like the Sun cannot erupt.

We shall also say a few words on the eruptions of supernovae. Applying to this case the formulae of the preceding section, we find that an amount of energy of the order of 10^{50} ergs is evolved in the eruptions of supernovae. This energy is comparable with the gravitational

energy of the star (if, of course, the mass of the supernova is not many times that of the Sun). Furthermore, the mass ejected in the eruption of a supernova is comparable with the mass of the star (again assuming that the mass of the supernova is of the order of the Sun's mass). Hence it may be supposed that, when a supernova erupts, a transition of the star takes place from one equilibrium configuration to another which differs considerably from the former. For example, it is not impossible that the eruption of a supernova signifies the transition of an ordinary star into the white dwarf state. However, supernova eruptions occur too rarely for all the white dwarfs in the Galaxy to be formed in this way in 10^{10} years.

PART VI.

STARS WITH BRIGHT SPECTRAL LINES

Chapter 28. The formation of emission lines

1. **Stars of the Wolf-Rayet, P Cygni and Be types.** This part will be devoted mainly to stars of early classes with bright spectral lines, i. e. stars of the Wolf-Rayet, P Cygni, and Be types. At present, the stars of these types have been studied more fully than those of other types having emission lines in their spectra. This is largely explained by the fact that the physical processes occurring in the atmospheres of the former stars are less complex.

Stars of the Wolf-Rayet type correspond to the spectral class O as regards the degree of excitation and ionisation of their atoms. Their spectra consist of broad bright bands of H, He I, He II, C III, N III and other atoms with very high ionisation potentials, superposed on a continuous background. Faint absorption lines are visible to the violet side of some of the bright bands. The width of the bright bands is some tens of ångströms, and the intensity within a band is sometimes 10 to 20 times greater than that of the continuous spectrum. In consequence of this, the energy emitted by a star in the bright lines is comparable with that emitted in the continuous spectrum (in the visible or the photographic region). Microphotometer tracings of the spectra of individual Wolf-Rayet stars are given in Fig. 66.

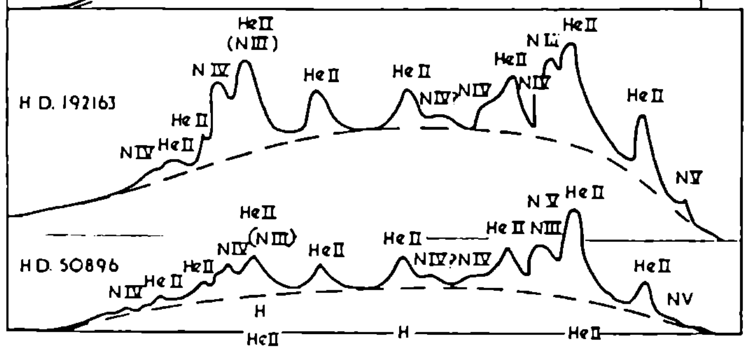

FIG. 66

An important property of the spectra of Wolf-Rayet stars is their division (admittedly not very rigorous) into two sequences, the *nitrogen* and the *carbon* sequences. In the spectra of the first sequence there are bands of nitrogen in various states of ionisation, but there are no bands of carbon or oxygen; in the spectra of the second sequence there are bands of carbon and oxygen in various states of ionisation, but none of nitrogen. The two sequences are similar as regards the degree of excitation and ionisation of the atoms. It seems that in this case there is an actual difference in the chemical composition of the stars.

In the last decade some spectroscopic binary systems with Wolf-Rayet-type components have been discovered. One of these systems (HD 193576) has been found to be an eclipsing variable also. The study of these systems has furnished much valuable information concerning Wolf-Rayet stars. In particular, it has been established that their masses are of the order of 10 solar masses.

The absolute magnitudes of Wolf-Rayet stars are of the order of -3^m. These stars are among the brightest objects in the Galaxy. However, as we know, spectra of the Wolf-Rayet type are also possessed by novae some years after eruption and by the nuclei of some planetary nebulae. These stars are considerably fainter than "ordinary" Wolf-Rayet stars. Their absolute magnitudes are on the average 3^m.

The stars of P Cygni type, which belong to the spectral class B, resemble the Wolf-Rayet stars. In the spectra of these stars, as in those of the Wolf-Rayet stars, bright lines are visible which lie approximately symmetrically relative to the central frequencies, and are bounded to the violet by absorption lines (Fig. 67a). However, the widths of the bright lines in this case are not so great as in the Wolf-Rayet spectra, while the absorption lines, on the other hand, are much more intense.

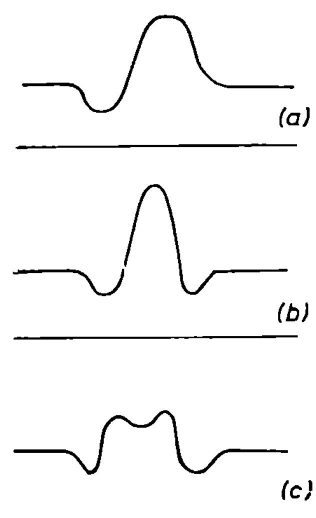

FIG. 67

Stars of P Cygni type are not the only B-type stars which have emission lines. A much greater number of these stars (called simply Be-type stars) have the line contours shown in Fig. 67 (b) and 67 (c). These contours can be described as follows: on a broad and shallow absorption line, a narrower emission line is superposed; the latter in some cases is single, but in others is double. The spectra of Be-type stars undergo considerable changes in the course of time. For example, the relative intensities of the components of the bright lines vary. Sometimes the bright lines disappear altogether, and the Be-type star becomes a normal

B-type star. As well as changes in the spectrum, small variations in the star's luminosity are observed.

2. **The outflow of matter from the stars.** The observational data on Wolf-Rayet and P Cygni stars lead to the conclusion that a continuous outflow of matter from these stars is taking place [18]. We shall enumerate the main facts which confirm this view.

Firstly, the line contours in the spectra of P Cygni and Wolf-Rayet type stars have the form shown in Fig. 67 (a). Such contours are most naturally explained by supposing that the lines are formed in envelopes ejected from the stars. Owing to the Doppler effect, the emission lines are broadened, and the absorption lines, which are formed in the part of the envelope approaching the observer, are displaced to the violet.

Secondly, lines in the spectra of novae have similar contours, and novae are objects which undoubtedly eject matter. However, unlike the spectra of novae, those of Wolf-Rayet and P Cygni type stars do not markedly change in the course of time. This compels us to suppose that we are here concerned not with the detachment and removal of envelopes, but with a continuous outflow of matter from the surfaces of the stars. Owing to this process, extended envelopes are formed around the stars, and the matter in them is continuously renewed.

Thirdly, if the broadening of the emission lines is caused by the Doppler effect, the width of the line should be $\Delta\lambda = 2\,v\,\lambda/c$, where v is the velocity of outflow of matter and c the velocity of light. This relation between the width $\Delta\lambda$ of the line and the wavelength λ is in fact satisfied for the spectra of Wolf-Rayet and P Cygni type stars. This makes it possible to determine the rate of outflow of matter from the stars. It is found to be of the order of 1000 km/sec for Wolf-Rayet stars and of the order of 100 km/sec for P Cygni type stars.

Fourthly, some Wolf-Rayet stars are components of close binaries. In such cases, to explain the spectroscopic and photometric phenomena, it is necessary to suppose that the companions of the Wolf-Rayet stars move in extended atmospheres which envelop these stars.

The existence of extended envelopes of Wolf-Rayet and P Cygni type stars makes it quite natural to assume that the emission lines in the spectra of these stars are formed as a result of fluorescence occurring in the same general way as in the planetary nebulae. As is well known, to bring about fluorescence the dilution factor must be small. This condition is in fact fulfilled in the extended envelopes. However, it must be noticed that the dilution factor in these envelopes is by no means so small as in the nebulae. We shall see below that, owing to this fact, the radiation processes in the extended envelopes are considerably more complicated than those in the nebulae.

It is very important to ascertain whether there is a noticeable acceleration or retardation of the atoms ejected from the stars. Some information on this point can be obtained from an analysis of the contours of spectral lines (see the next section). Another method of resolving this question is as follows. Since the radiation processes in these envelopes do not differ fundamentally from those in the planetary nebulae, there must exist in the envelopes the same stratification of the radiation as is found in the nebulae. In other words, the lines of atoms with high ionisation potentials must be formed in deeper layers of the envelopes than those of atoms with low ionisation potentials. For this reason, if a velocity gradient is present in the envelope, the emission lines of various atoms should be of different widths. It is evident that the widths of the lines will increase with the ionisation potential if the atoms are retarded in the envelope, and will decrease if they are accelerated. The observational data on the widths of emission lines of He I (λ 5876) and He II (λ 5411) in the spectra of some Wolf-Rayet stars are given in Table 23. It is seen from this table that the lines of neutral helium have greater widths than those of ionised helium. According to the interpretation we have just given, this indicates an accelerated motion of the atoms in the envelopes of Wolf-Rayet stars.

Table 23

Star		He I	He II
HD	192163	1815	1550
	177230	785	520
	192103	1290	975
	184738	1075	540

It is important also to estimate the amount of matter ejected by the star in unit time. It is evident that, in time dt, an amount of matter $\varrho(r)\, v(r)\, dt$ flows out through unit surface area at a distance r from the centre of the star, where $\varrho(r)$ and $v(r)$ are respectively the density and the velocity at the point in question in the envelope. The amount of matter flowing out through the whole spherical surface of radius r in time dt is

$$dM = 4\,\pi\, r^2\, \varrho(r)\, v(r)\, dt\,. \tag{28.1}$$

In order to estimate dM/dt, we take for r the radius of the photosphere, and for $\varrho(r)$ the value of the density at the inner boundary of the reversing layer, i. e. a density of the order of 10^{-11} g/cm^3 (this value of ϱ will be refined later). We find from formula (28.1) that a star of Wolf-Rayet or P Cygni type loses each year a mass equal to approxi-

mately 10^{-5} times the mass of the Sun. Such an immense quantity of mass lost annually by the star is evidence, above all, of the brevity of a star's existence in the Wolf-Rayet or P Cygni stage.

The idea that matter is ejected from Be stars and that they have extended envelopes is based both on the fact that emission lines are present in their spectra and on a number of peculiarities in these spectra. At present we shall mention only one of the peculiarities of Be star spectra. Judging from the contours of broad absorption lines, Be stars are among those which rotate rapidly. Their velocities of rotation at the equator reach some hundreds of kilometres per second. The matter ejected from the star, which forms the extended envelope, also takes part in the rotation, of course. This explains the broadening of the emission lines formed in the envelopes as a result of fluorescence. However, the rate of rotation of the ejected matter decreases as it moves away from the star (because of the conservation of angular momentum). For this reason the emission lines have smaller widths than the absorption lines on which they are superposed.

The rapid rotation of the Be stars apparently makes possible the ejection of matter from them. However, it cannot be supposed that the rotation itself is the cause of this ejection. This follows from the fact that the emission spectra of Be stars undergo irregular variations in the course of time (and sometimes disappear altogether). Consequently, the ejection of matter from Be stars is also irregular, which would be impossible if the matter were ejected as a result of the rotation.

The outflow of matter from Be stars is comparatively feeble. This is indicated by the fact that the presence of the envelope, except in particular cases, has little effect on the energy distribution in the continuous spectrum of the star. In the first approximation, the spectrum of a Be star can be regarded as the result of the superposition of emission lines on a more or less normal spectrum of class B. Consequently, the envelope of a Be star plays the part of the upper layers of the atmosphere. In this respect Be stars differ from Wolf-Rayet and P Cygni type stars, whose envelopes render the star itself completely invisible. The envelopes of stars of these latter types are so large that they serve not only as the atmospheres, but also as the photospheres of the stars. However, in speaking below about the envelopes of Wolf-Rayet stars, we shall have in mind only those parts in which the spectral lines are formed, i. e. the atmospheres. In Chapter 29 we shall discuss the envelopes of Wolf-Rayet stars as the photospheres in which the continuous spectrum is formed.

3. **The emission-line contours.** Valuable information on the nature of the motion of the matter ejected from a star can be obtained from a study of the emission-line contours in its spectrum.

In general, the theoretical determination of line contours formed by moving envelopes presents great difficulties. However, in the case where *the rate of motion of the envelopes is considerably greater than the mean thermal velocity of the atoms*, the line contours can be calculated fairly easily. This is due to the fact that, in this case, we can neglect all the factors affecting the contour of a line, except one: the motion of the envelope. In this case only, moreover, can the converse problem be solved with fair reliability: to find, from the observed line contours, the parameters which characterise the motion of the envelope. It may be supposed that this case occurs, to a very good approximation, in the envelopes of Wolf-Rayet, P Cygni and Be-type stars (and also of novae). For this reason we shall consider it in what follows.

Let us take a co-ordinate system XYZ, with the origin at the centre of the star and the z axis directed towards the observer. In order to determine the contour of an emission line, formed by the transition of an atom from the kth to the ith state, we calculate first the intensity of radiation $I_{ik}(x, y, \nu)$ coming from the point of the star's disc with co-ordinates x, y in a frequency ν within the line (Fig. 68). The total energy emitted in the frequency ν in the direction towards the observer (in unit solid angle) is then determined by the formula

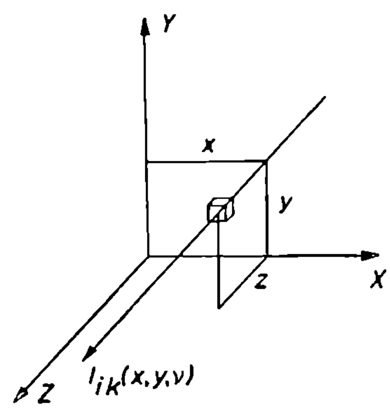

FIG. 68

$$\mathscr{E}_{ik}(\nu) = \iint I_{ik}(x, y, \nu)\, \mathrm{d}x\, \mathrm{d}y\ . \tag{28.2}$$

We shall not decide beforehand whether the envelope is transparent or opaque in the given line. Hence, besides the volume emission coefficient $\varepsilon_{ik}(\nu - \nu_{ik}')$, we introduce the volume absorption coefficient $\varkappa_{ik}(\nu - \nu_{ik}')$. Here ν_{ik}' denotes the central frequency of the line emitted by the element of volume concerned, which will in general be moving relative to the observer. This frequency is displaced, relative to the central frequency ν_{ik} of a line formed in a volume stationary with respect to the observer, by an amount

$$\nu_{ik}' - \nu_{ik} = \nu_{ik}\, v_z/c\ , \tag{28.3}$$

where v_z is the projection of the rate of motion of the element of volume considered on the z axis. We notice that the quantity ε_{ik} is proportional to the number of emitting atoms $n_k(x, y, z)$ in 1 cm^3, and the quantity $\varkappa_{ik}$ is proportional to the number of absorbing atoms $n_i(x, y, z)$ in 1 cm^3.

We have for the intensity of radiation $I_{ik}(x,y,\nu)$

$$I_{ik}(x,y,\nu) = \int_{-\infty}^{\infty} \varepsilon_{ik}(\nu - \nu_{ik}') \exp\left[-\int_{z}^{\infty} \alpha_{ik}(\nu - \nu_{ik}')\, dz\right] dz\,. \quad (28.4)$$

In writing formula (28.4), no assumptions have been made concerning the velocity field in the envelope. Now, in accordance with what was said above, let us assume that the rate of motion of the envelope considerably exceeds the thermal velocities of the atoms. In this case, without much loss of accuracy, we can replace ε_{ik} and α_{ik}, which are complicated functions of ν, by simpler ones. We shall suppose, in fact, that the quantities ε_{ik} and α_{ik} differ from zero, and are constant, in the frequency interval from $\nu_{ik}' - \frac{1}{2}\Delta\nu_{ik}$ to $\nu_{ik}' + \frac{1}{2}\Delta\nu_{ik}$, and are zero outside this interval, whilst

$$\Delta\nu_{ik} = 2\,\nu_{ik}\,u/c\,, \quad (28.5)$$

where u is the mean thermal velocity of the atoms.

It is clear that, on the assumptions made, the radiation of frequency ν will be emitted to the observer not by the whole envelope, but only by some region of it lying on both sides of the surface of equal radial velocities corresponding to the frequency ν. The equation of this surface has the form

$$\nu = \nu_{ik} + \nu_{ik}\, v_z(x,y,z)/c\,. \quad (28.6)$$

It is easy to see that the boundaries of the region mentioned will lie at a distance from the surface (28.6) along the line of sight (i. e. along the z axis) which corresponds to a change in frequency by $\frac{1}{2}\Delta\nu_{ik}$. Thus we obtain, instead of formula (28.4),

$$I_{ik}(x,y,\nu) = \int_{z_1}^{z_2} \varepsilon_{ik} \exp\left(-\int_{z}^{z_2} \alpha_{ik}\, dz\right) dz\,, \quad (28.7)$$

where the limits of integration z_1 and z_2 are determined from the conditions

$$\left.\begin{aligned} \nu - \tfrac{1}{2}\Delta\nu_{ik} &= \nu_{ik} + \nu_{ik}\, v_z(x,y,z_1)\,/\,c\,, \\ \nu + \tfrac{1}{2}\Delta\nu_{ik} &= \nu_{ik} + \nu_{ik}\, v_z(x,y,z_2)\,/\,c\,, \end{aligned}\right\} \quad (28.8)$$

or from conditions differing from (28.8) by the sign of $\frac{1}{2}\Delta\nu_{ik}$.

It is clear, moreover, that, on the assumptions made, the "thickness" of the layer which gives radiation in the frequency ν (i. e. the difference $z_2 - z_1$) will be comparatively small (except at particular points). This

enables us to assume that the values of α_{ik} and ε_{ik} are constant in this layer and equal to their values on the surface (28.6). Effecting the integration in formula (28.7), we find

$$I_{ik}(x, y, \nu) = (\varepsilon_{ik}/\alpha_{ik}) \left[1 - e^{-\alpha_{ik}(z_2 - z_1)}\right] . \tag{28.9}$$

The difference $z_2 - z_1$ can be found from the relations (28.8). Subtracting one of these relations from the other and taking

$$v_z(x, y, z_2) - v_z(x, y, z_1) = (z_2 - z_1)\,(\partial v_z/\partial z) , \tag{28.10}$$

we obtain

$$\Delta \nu_{ik} = \nu_{ik} \mid \partial v_z/\partial z \mid (z_2 - z_1) / c . \tag{28.11}$$

From (28.5), this gives

$$z_2 - z_1 = 2\,u / \mid \partial v_z/\partial z \mid . \tag{28.12}$$

We must now substitute the expression obtained for $I_{ik}(x, y, \nu)$ in formula (28.2), which determines the total energy $\mathcal{E}_{ik}(\nu)$ emitted by the envelope in the frequency ν. Taking account of (28.12), we find

$$\mathcal{E}_{ik}(\nu) = \iint \frac{\varepsilon_{ik}}{\alpha_{ik}} \left[1 - e^{-2u\alpha_{ik}/|\partial v_z/\partial z|}\right] \mathrm{d}x\,\mathrm{d}y . \tag{28.13}$$

Here the integration is extended over the surface (28.6).

The quantities ε_{ik} and α_{ik} which appear in formula (28.13) are expressed in the following well-known manner in terms of the concentrations of absorbing and emitting atoms n_i and n_k:

$$\varepsilon_{ik} = n_k\, A_{ki}\, h\nu_{ik}/4\,\pi\, \Delta\nu_{ik} , \tag{28.14}$$

$$\alpha_{ik} = \frac{n_i\, B_{ik}\, h\nu_{ik}}{c\, \Delta\nu_{ik}} \left(1 - \frac{g_i\, n_k}{g_k\, n_i}\right) , \tag{28.15}$$

where A_{ki} and B_{ik} are the Einstein transition coefficients. Taking into account the relation between A_{ki} and B_{ik}, we obtain

$$\frac{\varepsilon_{ik}}{\alpha_{ik}} = \frac{2\, h\nu_{ik}^3}{c^2}\, \frac{1}{(g_k\, n_i/g_i\, n_k) - 1} . \tag{28.16}$$

We notice that the relation (28.16) becomes Planck's formula, as it should, when n_k/n_i is determined by Boltzmann's formula.

Thus, to calculate the contour of an emission line, i. e. the quantity $\mathcal{E}_{ik}(\nu)$, it is necessary to know both the velocity distribution in the envelope and the distribution of absorbing and emitting atoms. In the

next section we shall show how the degree of excitation of the atoms in various parts of the envelope can be determined. The problem of calculating the contours of the emission lines will then be completely solved in principle. Here we shall give some further general considerations which hold for any envelopes.

First of all, it must be noted that the envelope can generally be divided into two regions: one opaque to radiation in the line concerned, and the other transparent to this radiation. In the first region the quantity $2\,u\,\alpha_{ik}/|\,\partial v_z/\partial z\,|$ exceeds unity, whilst in the second it is less than unity. According to this, the integral (28.13) which determines $\mathcal{E}_{ik}(\nu)$ can be approximately divided into two parts: firstly, the integral over that part of the surface (28.6) which lies in the opaque region:

$$\mathcal{E}_{ik}'(\nu) = \frac{2\,h\,\nu_{ik}^3}{c^2} \iint \frac{\mathrm{d}x\,\mathrm{d}y}{(g_k\,n_i/g_i\,n_k) - 1} \tag{28.17}$$

and, secondly, the integral over that part of the surface (28.6) which lies in the transparent region:

$$\mathcal{E}_{ik}''(\nu) = \frac{A_{ki}\,h\,\nu_{ik}}{4\,\pi\,\Delta\,\nu_{ik}} \iint n_k\,\frac{2\,u}{\partial v_z/\partial z}\,\mathrm{d}x\,\mathrm{d}y\,. \tag{28.18}$$

It is evident that for some frequencies inside the given line the quantity $\mathcal{E}_{ik}'(\nu)$ may predominate, and for others the quantity $\mathcal{E}_{ik}''(\nu)$.

The following question is of great interest. Is it possible, from observed line contours, to draw any conclusions about the total transparency of the envelope to radiation in these lines? It is easy to see that this can certainly be done if we have contours of two lines formed by transitions of an atom from the same upper level to two different lower levels. If the envelope is completely transparent to radiation in the lines concerned, it follows from formula (28.18) that the intensities within the lines, taken at corresponding points [i. e. for the same value of $(\nu - \nu_{ik})/\nu_{ik}$], will be proportional. On the other hand, if absorption of radiation in the lines takes place in the envelope, formula (28.17) shows that this proportionality will not hold. Consequently, from the fact that this proportionality is present, the conclusion follows that the envelope is completely transparent to radiation in the lines concerned.

However, it is not always possible to observe two lines with a common upper level, and it is found that this is not necessary. To solve the problem in question, it is sufficient to have contours of several lines of the same atom. If the envelope is completely transparent to radiation in the lines of the subordinate series of the atom concerned, the radiation of the envelope in the lines of this atom will be formed in the same way as the radiation of the nebulae in those lines, i. e. as a result of

photo-ionisations from the first state and subsequent recombinations. Here the ratio of the number of transitions in any line to the total number of recombinations will be the same at every point in the envelope. In other words, the quantity n_k can be represented in the form $n_k = z_k n_e n_+$, where z_k does not vary in the envelope. Owing to this circumstance, the intensities within the lines, for an envelope completely transparent in the lines of the subordinate series, will again satisfy the proportionality condition mentioned above.

Let us now consider one particular case which is important in applications. We assume that an outflow of matter from the star takes place with the same intensity in all directions. We denote by $v(r)$ the rate of motion of the matter at a distance r from the centre of the star, and by θ the angle between the direction of motion of the matter and the direction of the observer. Then the projected velocity in the line of sight is $v_z = v(r) \cos \theta$, and we easily obtain

$$\frac{\partial v_z}{\partial z} = \frac{\mathrm{d}v}{\mathrm{d}r} \cos^2 \theta + \frac{v}{r} \sin^2 \theta \,. \tag{28.19}$$

Instead of formula (28.13), we have in this case

$$\mathcal{E}_{ik}(\nu) = 2\pi \int \frac{\varepsilon_{ik}}{\alpha_{ik}} \left[1 - e^{-2u\alpha_{ik} / [(\mathrm{d}v/\mathrm{d}r)\cos^2\theta + (v/r)\sin^2\theta]}\right] a \, \mathrm{d}a \,, \tag{28.20}$$

where

$$a = r \sin \theta \,. \tag{28.21}$$

The integration in formula (28.20) is taken over the surface of equal radial velocities

$$\nu = \nu_{ik} + \nu_{ik} [v(r)/c] \cos \theta \,. \tag{28.22}$$

If the velocity of the matter does not vary within the envelope, formula (28.20) takes the form

$$\mathcal{E}_{ik}(\nu) = 2\pi \sin^2 \theta \int \frac{\varepsilon_{ik}}{\alpha_{ik}} \left(1 - e^{-2ur\alpha_{ik}/v\sin^2\theta}\right) r \, \mathrm{d}r \,, \tag{28.23}$$

where

$$\cos \theta = \frac{c}{v} \, \frac{\nu - \nu_{ik}}{\nu_{ik}} \,. \tag{28.24}$$

We find from formula (28.23) that, in the case of a completely transparent envelope,

$$\mathcal{E}_{ik}(\nu) = 4\pi (u/v) \int \varepsilon_{ik} r^2 \, \mathrm{d}r \,. \tag{28.25}$$

and in the case of a completely opaque envelope

$$\mathcal{E}_{ik}(\nu) = 2\,\pi \sin^2 \theta \int (\varepsilon_{ik}/\alpha_{ik})\, r\, \mathrm{d}r\,, \tag{28.26}$$

or, using (28.24),

$$\mathcal{E}_{ik}(\nu) = 2\,\pi \left[1 - \left(\frac{c}{v}\,\frac{\nu - \nu_{ik}}{\nu_{ik}}\right)^2\right] \int \frac{\varepsilon_{ik}}{\alpha_{ik}}\, r\, \mathrm{d}r\,. \tag{28.27}$$

Thus a completely transparent envelope gives an emission line with a rectangular contour (a) and a completely opaque one gives an emission line with a parabolic contour (b). Both these contours are shown in Fig. 69.

It may be thought that a continuous ejection of matter with the same intensity in all directions takes place from Wolf-Rayet type stars. A detailed study of emission-line contours in the spectra of these stars has shown that in one case (the star HD 193793) the line contours have a rectangular form. This indicates the complete transparency of the envelope to radiation in the lines, and the absence of an appreciable velocity gradient in the envelope. Usually, however, the emission-line contours are rounded; this can be explained either by an acceleration or retardation of the atoms ejected from the star, or by the opacity of the envelope to radiation in the lines. A comparison of the contours of the ionised helium lines λ 4686, 5411 and 4339 Å in the spectrum of the star HD 192163 has shown that the ordinates of the contours at corresponding points are proportional; consequently, in this case we must take the former of the two possibilities mentioned. For this star the velocity distribution in the envelope has been calculated from the form of the contour.

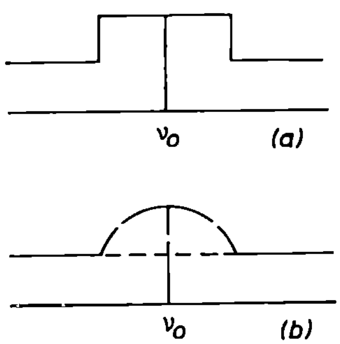

Fig. 69

However, for some lines the envelopes of Wolf-Rayet stars are certainly opaque. Among these are, first of all, lines whose lower state is metastable. As an example we may give the line $\lambda = 3889$ Å of neutral helium. Its lower level 2^3S has, as is well known, an extremely long lifetime. The great optical thickness of the envelopes of Wolf-Rayet stars in the line 3889 Å can be seen from the fact that, in the spectra of all Wolf-Rayet stars, the corresponding emission band is bounded on the violet side by a broad and deep absorption line.

Using formula (28.13), we can determine the emission-line contours in the case where the ejection of matter takes place from rapidly rotating stars. It is found that, in this case, the emission line contours have the

same form as those of lines in the spectra of Be-type stars (see Fig. 67 b, 67 c). These calculations confirm the view that the Be-type stars are rapidly rotating stars from which matter is being ejected. The irregular variations, observed in the spectra of these stars, in the intensities and shapes of the spectral lines are then explained by variations of density and velocity in the envelope, i. e. irregularities in the ejection of matter from the star.

Let us calculate the total energy emitted by the envelope in a given line. To do so, we must integrate the quantity $\mathcal{E}_{ik}(\nu)$ over all frequencies and directions, i. e. calculate the integral

$$E_{ik} = \iint \mathcal{E}_{ik}(\nu)\, \mathrm{d}\nu\, \mathrm{d}\omega\ . \tag{28.28}$$

Using the expression (28.13) for $\mathcal{E}_{ik}$, and changing from ν to z as variable of integration by means of the relation (28.6), we obtain

$$E_{ik} = 4\,\pi\, \Delta\nu_{ik} \iiint \varepsilon_{ik}\, \beta_{ik}\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z\ , \tag{28.29}$$

where we have put

$$\beta_{ik} = \int \left(1 - e^{-2\,u\,\alpha_{ik}/\left|\partial v_z/\partial z\right|}\right) \frac{1}{2\,u\,\alpha_{ik}} \left|\frac{\partial v_z}{\partial z}\right| \frac{\mathrm{d}\omega}{4\,\pi}\,. \tag{28.30}$$

The quantity β_{ik} has a simple physical significance. Since $4\,\pi\,\Delta\nu_{ik}\,\varepsilon_{ik}$ is the total energy emitted in the given line by unit volume, β_{ik} is simply the fraction of this energy which emerges from the envelope. If the envelope is opaque in all directions at the point concerned,

$$\beta_{ik} = \frac{1}{2\,u\,\alpha_{ik}} \left|\frac{\partial v_z}{\partial z}\right|_{\mathrm{av}}\,, \tag{28.31}$$

where $\left|\partial v_z/\partial z\right|_{\mathrm{av}}$ is the value of $\left|\partial v_z/\partial z\right|$ averaged over all directions. If the envelope is transparent in all directions at the point concerned (i. e. $2\,u\,\alpha_{ik}/\left|\partial v_z/\partial z\right| \ll 1$), then

$$\beta_{ik} = 1\ , \tag{28.32}$$

as it should. In the next section we shall discuss formula (28.30).

4. **The emission-line intensities.** We have previously calculated the intensities of the emission lines for the planetary nebulae. There it was assumed that ionisation of the atoms from the ground state takes place in the nebulae under the action of the radiation from the central star; then there follow captures of the electrons by the ions, and then transitions of the electrons from higher to lower levels, accompanied by the emission of quanta which leave the nebula unhindered (except for

quanta of the principal series). This latter circumstance, i. e. the complete transparency of the nebula to radiation in the lines of the subordinate series, considerably facilitated the determination of the emission line intensities. We recall that this determination reduced to the solution of a system of linear algebraic equations (expressing the steady-state condition for each energy level of the atom concerned).

Let us now consider the calculation of the intensities of bright lines formed in the envelopes of Wolf-Rayet, P Cygni and Be-type stars (and also of novae). The radiation of these envelopes is produced in fundamentally the same way as that of the nebulae. However, the degree of excitation of atoms in these envelopes is by no means so low as in the nebulae. For this reason, ionisation may occur in them not only from the ground state, but also from excited states, and the envelopes may be opaque to radiation in the lines of the subordinate series. In such conditions, the quanta formed in the central parts of the envelope (after ionisation and recombination) can leave the envelope only by coming close to its outer boundary, having undergone on the way a large number of processes of scattering, partition, etc. It is therefore clear that the problem of calculating the intensities of the bright lines reduces in this case to the solution of a very complex system of integro-differential equations, consisting of both the conditions of the steady state for each level and the equations of radiative transfer for each line and each continuum. This system of equations was first considered by V. A. AMBARTSUMYAN [8].

If the motion of the envelope is taken into account, the problem is still further complicated. However, if the velocity gradient in the envelope is fairly large (as is in fact the case in the stellar envelopes concerned), the theory is simplified. When a velocity gradient is present in the envelope, the quanta in the lines can emerge not only from its boundary regions, but also (owing to the Doppler effect) from its internal regions. For a fairly large velocity gradient, therefore, the solution of the problem of the radiation in a line at some point in the envelope will depend only slightly on what is radiated in that line at other points in the envelope. If this is true, the problem of finding the bright-line intensities again reduces approximately to the solution of some system of algebraic (though not linear) equations [154].

Let us consider this problem in more detail. Let $n_i B_{ic} \varrho_{ic}$ be the number of ionisations from the ith level, $n_e n_+ C_{if}(T_e)$ the number of captures into the ith level, and $n_k A_{ki}$ the number of spontaneous transitions from the kth to the ith level, taking place at some point in the envelope in 1 cm³ in 1 second.

If there were no velocity gradient in the envelope, the transitions from the kth to the ith level taking place in the central parts of the

envelope would be balanced by the transitions from the ith to the kth level, since almost all quanta emitted in the spectral lines would be absorbed in the envelope itself. However, if a velocity gradient is present, the number of transitions from the kth to the ith level will be greater than the number of converse transitions, since some fraction of the quanta in the line concerned leave the envelope because of the Doppler effect. We denote this fraction by β_{ik}. Then the excess of the number of transitions $k \to i$ over the number of converse transitions is $n_k A_{ki} \beta_{ik}$.

In the steady state, the number of transitions of atoms from the ith state to all others must equal the number of transitions to the ith state. Hence we obtain

$$n_i \left(\sum_{k=1}^{i-1} A_{ik} \beta_{ki} + B_{ic} \varrho_{ic}\right) = \sum_{k=i+1}^{\infty} n_k A_{ki} \beta_{ik} + n_e n_+ C_{if}(T_e) . \quad (28.33)$$

In these equations, the quantities ϱ_{ic} are supposed known and equal to

$$\varrho_{ic} = W \varrho_{ic}^*, \quad (28.34)$$

where ϱ_{ic}^* is the density of radiation beyond the limit of the ith series at the surface of the star, and W is the dilution factor.

The quantities β_{ik} have already been calculated in the preceding section, in determining the contours of emission lines formed by moving envelopes. However, we prefer not to use these results, but to calculate the quantities β_{ik} afresh, by a direct method. Here, as before, we shall suppose that both the absorption coefficient α_{ik} and the emission coefficient ε_{ik}, in a line of frequency ν_{ik}, are constant and non-zero in an interval $\Delta \nu_{ik} = 2 u \nu_{ik}/c$, and are zero outside this interval. Furthermore, weassume that the region of the envelope in which radiation in the given line is absorbed is comparatively small (because of the large velocity gradient), so that the density of matter and the velocity gradient in this region can be regarded as constant.

Let us consider the fate of quanta in a line of frequency ν_{ik} which are emitted by some element of volume in the z direction within a solid angle $d\omega$. It is evident that between z and $z + dz$ a fraction

$$e^{-\alpha_{ik} z} \left(1 - \frac{\nu'_{ik} - \nu_{ik}}{\Delta \nu_{ik}}\right) \alpha_{ik} \, dz \quad (28.35)$$

of these quanta will be absorbed. Here the factor $e^{-\alpha_{ik} z}$ takes account of the absorption of quanta between 0 and z, and the factor

$$1 - (\,|\,\nu'_{ik} - \nu_{ik}\,|/\Delta \nu_{ik})$$

takes account of the change in frequency of the radiation as a result of the Doppler effect;

$$\nu'_{ik} - \nu_{ik} = (\nu_{ik} z/c)\, \partial v_z/\partial z\,. \qquad (28.36)$$

It is clear, moreover, that the radiation in the given line will be absorbed only between 0 and z_1, where z_1 is determined from the condition

$$\Delta\nu_{ik} = (\nu_{ik} z_1/c)\, |\, \partial v_z/\partial z\, |\,. \qquad (28.37)$$

Using (28.36) and (28.37), the expression (28.35) takes the form

$$e^{-\alpha_{ik} z}(1 - z/z_1)\, \alpha_{ik}\, \mathrm{d}z\,. \qquad (28.38)$$

Multiplying the expression (28.38) by $1/4\pi$ and integrating it, first over z from 0 to z_1 and then over the solid angle, we obtain the fraction that are absorbed in the envelope out of the total number of quanta emitted in the given volume. With the above notation, this fraction is $1 - \beta_{ik}$. Hence we find for β_{ik}

$$\beta_{ik} = \int (1 - e^{-\alpha_{ik} z_1})\, \frac{1}{\alpha_{ik} z_1}\, \frac{\mathrm{d}\omega}{4\pi}\,. \qquad (28.39)$$

If we substitute the value of z_1 determined by the condition (28.37) in formula (28.39), we again reach the formula (28.30) previously obtained.

Thus, to determine the degree of excitation and ionisation of the atoms in the envelope, we have the system of equations (28.33) together with the relations (28.30). The absorption coefficients α_{ik} which appear in these relations are determined by formula (28.15), given in the preceding section.

If the equations (28.33) are solved for various parts of the envelope, we are enabled to determine the total amount of energy emitted by the envelope in any spectral line. The following obvious formula serves for this purpose:

$$E_{ik} = A_{ki}\, h\nu_{ik} \int n_k\, \beta_{ik}\, \mathrm{d}V\,, \qquad (28.40)$$

where the integration is extended over the whole volume of the envelope. It is easy to see that formula (28.40) is the same as formula (28.29) obtained in the preceding section.

Since the conditions are not the same at different points in the envelope, the equations (28.33) will have different forms at such points. The conditions are simplest in the outermost parts of the envelope, where they are similar to those in the nebulae. We can suppose that these parts of the envelope are completely transparent to radiation in the lines of the subordinate series, and hence we can put in equations (28.33)

$$\beta_{ik} = 1 \quad (i = 2, 3, 4, \ldots)\,.$$

Neglecting also the ionisation from excited states, we obtain instead of equations (28.33)

$$n_i(A_{i1}\beta_{1i}+\sum_{k=2}^{i-1}A_{ik})=\sum_{k=i+1}^{\infty}n_k A_{ki}+n_e n_+ C_{if}(T_e) \quad (i=2,3,4,\ldots). \quad (28.41)$$

In the lines of the principal series, the envelope is highly opaque, i. e. $\beta_{1i} \ll 1$. As is seen from equations (28.41), we can assume $\beta_{1i}=0$ for all levels except the second, without great error. Hence we have for the outer parts of the envelope

$$n_i \sum_{k=2}^{i-1} A_{ik}=\sum_{k=i+1}^{\infty} n_k A_{ki}+n_e n_+ C_{if}(T_e) \quad (i=3,4,5,\ldots). \quad (28.42)$$

The equations (28.42) have already been considered in Part IV. The solution of these equations was given there for the hydrogen atom, and the Balmer decrement was calculated. We recall that it was found to depend very little on the electron temperature ($H_\alpha/H_\beta=3$, $H_\gamma/H_\beta=0{\cdot}5$).

Another limiting case is obtained for those parts of the envelope where it is opaque to radiation in the lines of all series. In this case we can take the expression (28.31) for the value of β_{ik}. Substituting this expression in the equations (28.33) and using the fact that

$$A_{ki}\,\beta_{ik}=\frac{(g_2/g_1)\,n_1-n_2}{(g_k/g_i)\,n_i-n_k}\left(\frac{\nu_{ki}}{\nu_{12}}\right)^3 A_{21}\,\beta_{12}\,, \quad (28.43)$$

we obtain

$$n_i\left[x\sum_{k=1}^{i-1}\frac{(g_2/g_1)\,n_1-n_2}{(g_i/g_k)\,n_k-n_i}\left(\frac{\nu_{ki}}{\nu_{12}}\right)^3+\frac{B_{ic}\,\varrho_{ic}^*}{A_{21}}\right]=$$

$$x\sum_{k=i+1}^{\infty} n_k\,\frac{(g_2/g_1)\,n_1-n_2}{(g_k/g_i)\,n_i-n_k}\left(\frac{\nu_{ik}}{\nu_{12}}\right)^3+\frac{n_e n_+ C_{if}}{WA_{21}}\,, \quad (28.44)$$

where we have put

$$x=\beta_{12}/W\,. \quad (28.45)$$

We have arrived at a system of algebraic equations in the unknowns n_i/n_1 and $n_e\,n_+/Wn_1$. The parameters appearing in these equations are, besides the quantity x, the temperature of the star (which enters through $B_{ic}\,\varrho_{ic}^*$) and the electron temperature of the envelope (which enters through C_{if}). The system of equations (28.44) can easily be solved numerically. As an example, we give below the Balmer decrement calculated from the solution of this system for the hydrogen atom (for the case $T_*=20{,}000°$, $T_e=20{,}000°$).

Table 24

x	0	0·01	0·1	1·0
H_α	0·67	0·98	2·0	5·20
H_β	1·00	1·00	1·00	1·00
H_γ	0·97	0·79	0·44	0·21
H_δ	0·87	0·58	0·22	0·06

Let us turn now to the observational data. The Balmer decrement in the spectra of Wolf-Rayet type stars can be determined, though with some difficulty, since the hydrogen lines in these spectra coalesce with those of ionised helium. Some difficulties arise also in finding the Balmer decrement in the spectra of Be-type stars, because in this case the bright lines are superposed on absorption lines. Nevertheless, it may be asserted that the Balmer decrement in the spectra of Be-type stars differs markedly from that in the spectra of the planetary nebulae. For example, the ratio of intensities of the H_α and H_β lines in the spectra of different Be stars is both greater than 3 and less than 3 (on the average, for a group of stars, $H_\alpha/H_\beta = 2{\cdot}25$). Moreover, the Balmer decrement varies not only from one star to another, but also in the spectrum of a single star in the course of time.

The explanation of this behaviour of the Balmer decrement in the spectra of Be stars consists in the fact that the envelopes of these stars have not only parts transparent to radiation in the lines of the subordinate series, i. e. similar to the conditions of radiation in a nebula, but also parts not transparent to radiation in the lines of some subordinate series. The radiation in the lines which reaches the observer is composed of the radiation from the various parts of the envelope (in proportions which vary in the course of time).

The Balmer decrement is similar in the spectra of novae at times not very far from the moment of maximum luminosity. For example, in the spectrum of Nova Herculis 1934 during the first three months after the outburst, the ratio of intensities of the H_α and H_β lines was on the average 1·9.

It is interesting to notice that, in the spectra of some novae several months after the outburst, a very large ratio of intensities of the H_α and H_β lines has been observed. Thus, in the spectrum of Nova Lacertae 1936 this ratio was between 5 and 6, and in that of RS Ophiuchi 1933 it reached 10 or 12. This phenomenon is apparently explained by the fact that, in the periods when the observations were made, considerable parts of the envelopes of these stars were not transparent to radiation in the lines of the Lyman and Balmer series, but were transparent

to radiation in the lines of other series. It is found that in this case the solution of equations (28.33) does in fact lead to such large values of H_α/H_β.

From a detailed comparison of the theory explained above with observation, we can find a number of important characteristics of stellar envelopes. As an example, let us compare the theoretical and observed values of the line intensities of hydrogen and ionised helium in the spectra of Wolf-Rayet stars. Here, for simplicity (and also taking account of the inexactness of the observational results), we shall suppose that the envelopes of Wolf-Rayet stars are completely transparent in the lines of the subordinate series of these atoms.

If the envelope is transparent to radiation in the lines of the subordinate series of a given atom, the total energy emitted by the envelope in any of these lines, according to formula (28.40) (with $\beta_{ik} = 1$), is

$$E_{ik} = 4\,\pi\, A_{ki}\, h\nu_{ik} \int\limits_{r_0}^{r_1} n_k\, r^2\, \mathrm{d}r\,, \tag{28.46}$$

where r_0 is the radius of the inner boundary of the envelope and r_1 the radius of the outer boundary of that part of the envelope which radiates in the lines of the atom concerned. In this case the equations (28.42) serve to determine the values of n_k. From them we can find the values of $z_k = n_k/n_e\, n_+$, which depend only on T_e. Assuming that T_e does not vary in the envelope, we obtain instead of formula (28.46)

$$E_{ik} = 4\,\pi\, A_{ki}\, h\nu_{ik}\, z_k \int\limits_{r_0}^{r_1} n_e\, n_+\, r^2\, \mathrm{d}r\,. \tag{28.47}$$

In calculating the integral (28.47), we assume that the atoms ejected from the envelope move with constant velocity. Then, according to the relation (28.1), the density in the envelope decreases inversely as the square of the radius, and we can write

$$n_e = n_e^0 (r_0/r)^2\,, \quad n_+ = n_+^0 (r_0/r)^2\,, \tag{28.48}$$

where n_e^0 and n_+^0 are the values of n_e and n_+ at the inner boundary of the envelope. Substituting the expressions (28.48) in formula (28.47), and taking for simplicity $r_1 = \infty$, we find

$$E_{ik} = 4\,\pi\, A_{ki}\, h\nu_{ik}\, z_k\, n_e^0\, n_+^0\, r_0^3\,. \tag{28.49}$$

A very interesting result is obtained on applying formula (28.49) to determine the relative content of hydrogen and helium in the atmospheres of Wolf-Rayet stars [4]. Applying formula (28.49) to the line

$\lambda = 4686$ Å of ionised helium (transition $4 \to 3$) and to the H_β line of hydrogen, we obtain from these two formulae

$$\frac{E_{34}'}{E_{24}} = \frac{A_{43}'\, \nu_{34}'\, z_4'\, n_+^{0\prime}}{A_{42}\, \nu_{24}\, z_4\, n_+^{0}}, \tag{28.50}$$

where the primed quantities refer to ionised helium, and the others to hydrogen (that is, $n_+^{0\prime}$ is the number of doubly ionised helium atoms in 1 cm^3 at the inner boundary of the envelope). For the star HD 192163, the ratio of intensities of the lines $\lambda = 4686$ Å and H_β is approximately 20 [18]. We also have $A_{43}'/A_{42} = 16$, $\nu_{34}'/\nu_{24} \approx 1$, $z_4'/z_4 \approx \frac{1}{2}$ (for $T_e = 50{,}000°$). With these data, we find from formula (28.50) $n_+^{0\prime}/n_+^0 = 2{\cdot}5$. Thus we reach the conclusion that, in the atmospheres of Wolf-Rayet stars, the number of helium atoms is approximately 2·5 times the number of hydrogen atoms. This result deserves considerable attention because, in the atmospheres of other stars and in the planetary nebulae, the ratio of the numbers of hydrogen and helium atoms is about 10. Consequently, the atmospheres of Wolf-Rayet stars have a chemical composition which differs from that of the atmospheres of other stars. This fact is undoubtedly very important for an understanding of the nature of Wolf-Rayet stars, especially when we recall that, according to present-day views, the source of stellar energy is formed by nuclear reactions which convert hydrogen into helium.

However, it must be remarked that the estimate made above of the relative abundance of hydrogen and helium in the atmospheres of Wolf-Rayet stars is rather rough, since it does not take account of the possible opacity of the envelopes of Wolf-Rayet stars to radiation in the lines of the subordinate series of the atoms concerned, or of certain other factors.

Using formula (28.49), we can also estimate the concentration of free electrons at the inner boundary of the envelope (i. e. at the surface of the "photosphere"). Taking into account the facts that the most abundant element in the atmospheres of Wolf-Rayet stars is helium, and that the helium atoms are predominantly in the doubly ionised state, we can take $n_e = 2\, n_+'$. Hence we obtain, on applying formula (28.49) to the line λ 4686 Å of ionised helium,

$$E_{34}' = 4\,\pi\, A_{43}'\, h\nu_{34}'\, z_4'\, \tfrac{1}{2}\, (n_e^0)^2\, r_0^3\,. \tag{28.51}$$

On the other hand, the energy emitted by the envelope in the line $\lambda = 4686$ Å can be expressed in terms of the quantity A_{4686} found from observation:

$$E_{34}' = \frac{8\,\pi^2\, r_0^2\, h\, (\nu_{34}')^4}{c^2}\, \frac{1}{e^{h\nu_{34}'/kT_*} - 1}\, A_{4686}\,. \tag{28.52}$$

The relations (28.51) and (28.52) give

$$A_{43}' \, z_4' \, (n_e^0)^2 \, r_0 = \frac{4 \pi (\nu_{34}')^3}{c^2} \, \frac{1}{e^{h\nu_{34}'/kT_*} - 1} \, A_{4686} \, . \tag{28.53}$$

To determine n_e^0 from formula (28.53), we must know from observation, besides the value of A_{4686}, the radius r_0 of the photosphere and the temperature T_* of the star. Unfortunately these quantities are imperfectly known for Wolf-Rayet stars. However, since n_e^0 enters as a square in formula (28.53), the exact values of these quantities are unnecessary for an estimate of n_e^0. We shall take $r_0 = 5 \, r_\odot$, $T_* = 50{,}000°$. For Wolf-Rayet stars, the quantities A_{4686} have values of the order of some hundredths (for example, for the star HD 192163 $A_{4686} = 0{\cdot}047$). With these values, we obtain from formula (28.53) $n_e^0 = 1{\cdot}5 \times 10^{12}$. This value has already been used above in determining the amount of matter ejected each year by a Wolf-Rayet star.

Chapter 29. Problems of the physics of stars with bright spectral lines

1. **The temperatures of the stars.** To determine the temperatures of stars with bright lines in their spectra, we can use the method proposed by H. ZANSTRA for determining the temperatures of the nuclei of planetary nebulae. This method is based on the hypothesis that the radiation of the nebula in the lines of any atom is produced at the expense of the energy from the star beyond the limit of the principal series of this atom. To determine the temperature of the star, we have the equation

$$\int_{x_0}^{\infty} \frac{x^2 \, dx}{e^x - 1} = \sum \frac{x_i^3 A_i}{e^{x_i} - 1} \, , \tag{29.1}$$

where $x_0 = h\nu_0/kT_*$, $x_i = h\nu_i/kT_*$, ν_0 is the ionisation frequency of the atom, ν_i is the frequency of an observed line, and $\nu_i A_i$ is the ratio, obtained from observation, of the energy emitted by the nebula in the line to that emitted by the star in unit frequency interval of the continuous spectrum near the line.

This method has several times been applied to determine the temperatures of stars of the Wolf-Rayet, P Cygni, and Be types. Recently, B. A. VORONTSOV-VEL'YAMINOV [172] has determined the temperatures of six Wolf-Rayet type stars, and C. S. BEALS and R. D. HATCHER [19] those of 52 stars, of spectral classes A 4 to O 6, having emission lines in their spectra.

The results of B. A. Vorontsov-Vel'yaminov are given in Table 25. The first column gives the number of the star in the Henry Draper catalogue, and the other columns give the temperatures of the stars (in thousands of degrees) found from the lines of various ions (the ionisation potentials in electron-volts are shown below the symbols of the ions). The last column will be explained later.

Table 25

Star	He I 24·5	C III or N III 47·4	He II 54·2	C IV 64·2	N IV 77·0	T_c
HD 192163	32	65	73		84	15
191765	35	62	69		75	15
193077	29	51	59		74	13
193576	29	48	60		62	14
192103	33	64	63	69		12
192641		59	55	70		7

The results of Beals and Hatcher are shown in Fig. 70. On the axis of abscissae are placed the spectral classes, on the axis of ordinates the temperatures of the stars, as found from lines of H, He I, Fe II, He II, N III and C III. The continuous line shows the dependence of temperature on spectral class which, in the opinion of these authors, is the most probable.

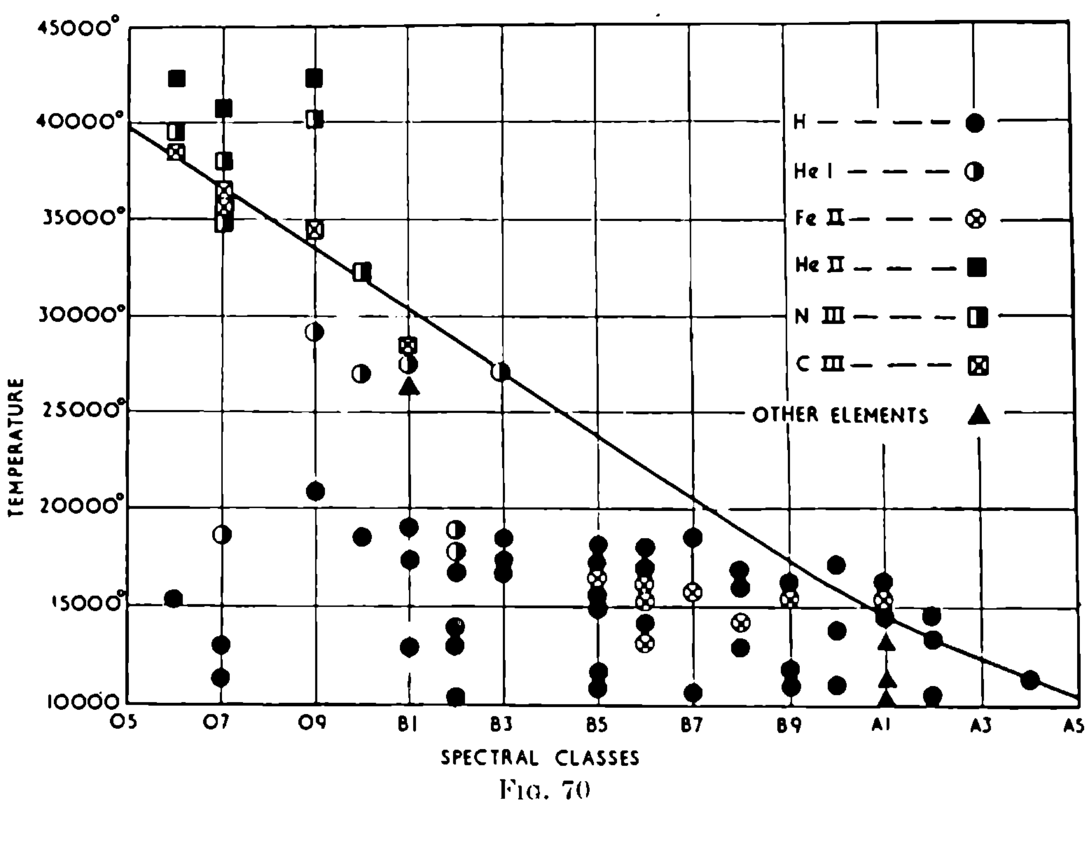

Fig. 70

It is seen from the observational data that the temperature of the same star, as found from the lines of various atoms, has very different values. Here it is immediately obvious that the temperature increases with the ionisation potential of the atom from whose lines it is determined. It is possible that this difference is to some extent explained by the deviation of the star's radiation from Planck's Law. On the other hand, there are a number of factors which bring about large errors in the temperatures of the stars when determined by this method. We shall enumerate the most important of these factors.

(1) In the derivation of equation (29.1) it was assumed that the envelope absorbs all the radiation from the star beyond the limit of the principal series of the atom, i. e. that its optical thickness beyond this limit is greater than unity. In general we have, instead of equation (29.1),

$$\int_{x_0}^{\infty} \frac{x^2}{e^x - 1} [1 - e^{-\tau_0(x_0/x)^3}] \, dx = \sum \frac{x_i^3 A_i}{e^{x_i} - 1}, \tag{29.2}$$

where τ_0 denotes the optical thickness of the envelope immediately beyond the limit of the principal series of the atom, and it is assumed that the absorption coefficient is inversely proportional to the cube of the frequency ($k_\nu \sim 1/\nu^3$). If $\tau_0 \ll 1$, the use of equation (29.1) instead of equation (29.2) will clearly lead to too low a temperature for the star. This effect should be important for atoms with comparatively low ionisation potentials, in consequence of their high ionisation in the atmospheres of hot stars and the resulting small value of τ_0. For example, the optical thicknesses of the atmospheres of Wolf-Rayet stars, beyond the limit of the Lyman series, are much less than unity. As a result, their temperatures as found from hydrogen lines are of the order of 20,000°, which is very low for Wolf-Rayet stars.

(2) Strictly speaking, the summation on the right-hand side of equation (29.1) should be extended over all the lines in the second series of the atom (including the continuum). In practice, however, one line, or at most a few lines, of the second or other series are observed. Hence, to find the exact value of the temperature, it is necessary to know the ratio of the number of quanta emitted by the envelope in the second series to the number emitted in the observed line. Since the number of quanta emitted in the second series is equal to the number of captures into all levels from the second upwards, this ratio is, for lines formed by the transition from the kth to the ith level,

$$q - n_e n_+ [\sum_{i=2}^{\infty} C_{if}]/n_k A_{ki} \, . \tag{29.3}$$

The ratio q can be calculated by means of the quantities $z_k = n_k/n_e\, n_+$. determined from the equations of the steady state for the energy levels of the atom. For example, we find for the H_β line of hydrogen that q is approximately equal to 9 (depending very slightly on the electron temperature T_e). Almost the same value of q is obtained for the brightest of the observed lines of ionised helium, $\lambda = 4686$ Å, since the H and He II atoms are similar, and the transition probabilities A_{42} and A_{43} are nearly the same.

If the ratio q is known, the temperature of the star can be determined from the equation

$$\int_{x_0}^{\infty} \frac{x^2\,\mathrm{d}x}{e^x - 1} = q\,\frac{x_1{}^3}{e^{x_1} - 1}\,A_1\,, \tag{29.4}$$

where x_1 and A_1 refer to the line concerned. However, the values of q are not known at present for all atoms. Hence the factor q in equation (29.4) is usually either omitted or chosen arbitrarily. The error arising from this may be considerable. Thus, for instance, taking $q = 2$ for the line $\lambda = 4686$ Å of ionised helium, instead of the correct value $q = 9$, we obtain a temperature of 60,000° instead of the more correct value of 70,000°.

(3) In the derivation of equation (29.1), it was assumed that the envelope radiates in the lines of the atom concerned only at the expense of the energy from the star beyond the limit of the principal series of this atom. However, in some cases the radiation of the envelope in the lines of one atom may be excited by the radiation of the envelope in the lines of another atom, as a result of the chance proximity of the frequencies of particular lines of these atoms. Owing to the presence of a velocity gradient in the envelope, the probability of such processes is considerably increased. As an example, we may indicate the possibility of transitions of He II ions from the second state to the fourth under the action of the radiation of hydrogen atoms in the Lyman α line (the corresponding wavelengths are $\lambda = 1215{\cdot}18$ Å and $\lambda = 1214{\cdot}68$ Å). Later, the He II ion may spontaneously pass from the fourth state to the third, giving the observed line $\lambda = 4686$ Å (Fig. 71). Another such example is the possible excitation, by radiation in the resonance line

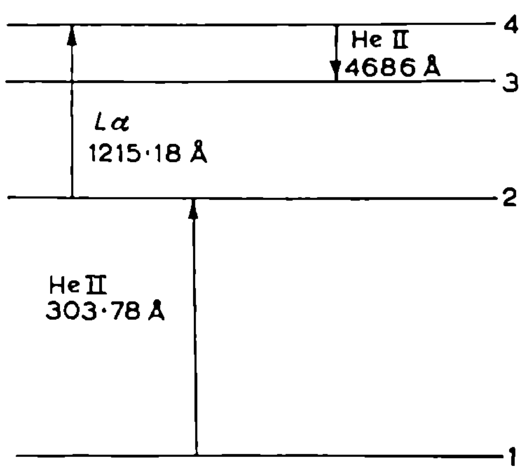

FIG. 71

of He II $\lambda = 304$ Å, of one of the levels of the O III ion; the cascade transitions of the electrons from this level are accompanied by the emission of quanta in the visible part of the spectrum. It is of interest that the return of the electron to its previous level in the O III ion is accompanied by the emission of a quantum in the line $\lambda = 374$ Å, which in turn may excite one of the levels of the N III ion, thereby bringing about the appearance of a number of observed lines of this ion (Fig. 72).

If these factors are in operation, the temperatures of the stars, as determined from the lines of He II and N III, will evidently be too high. BEALS and HATCHER think that this is indeed the case for Wolf-

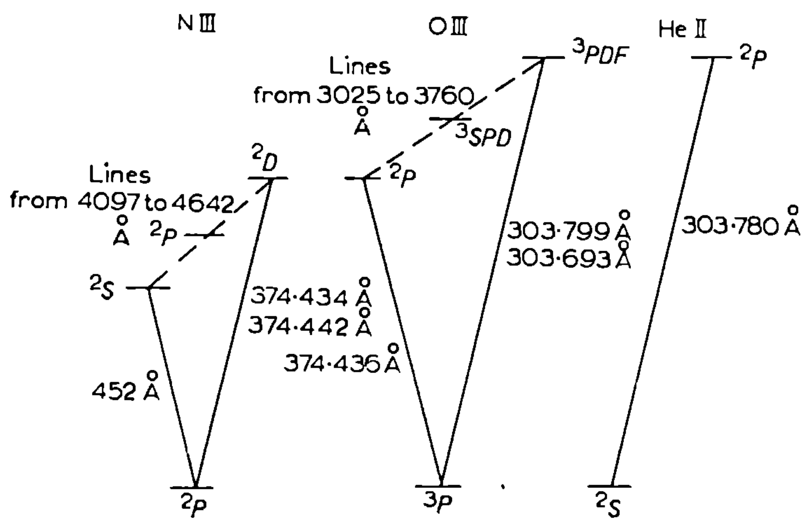

FIG. 72

Rayet stars, an opinion which is reflected in the curve given by them in Fig. 70. The temperatures found from the lines of other atoms are predominantly too low, in consequence of the effect of the first factor mentioned (the incomplete absorption by the envelope of the energy from the star beyond the limit of the principal series of the atom).

However, even when the factors enumerated above have been taken into account, we cannot be sure that it is sufficiently justified to apply this method to determine the temperatures of Wolf-Rayet, P Cygni and Be stars (and of novae in their early stages). As we know, the envelopes of these stars have comparatively small radii, and the dilution factor in them is much greater than in the nebulae. For this reason, the radiation of these envelopes is due to more complex processes than that of the nebulae. Bearing in mind the application of this method to the problem of determining the temperatures of stars, we shall now consider more closely the radiation of envelopes of small radius.

Let us assume that the envelope consists of atoms having only three energy levels. Under the action of the radiation from the star, the atoms make both cyclic transitions of the kind $1 \to 3 \to 2 \to 1$ and the converse transitions of the kind $1 \to 2 \to 3 \to 1$. In discussing the nebulae, it was shown that the number of transitions of the first kind is approximately $1/W$ times the number of transitions of the second kind. However, the diffuse radiation of the envelope itself was not taken into account. Nevertheless, the density of the diffuse radiation in the frequency ν_{12} (i. e. in the Lyman α line, if we consider the hydrogen atom) is very high. For this reason, the number of transitions of the kind $1 \to 2 \to 3 \to 1$ is greatly increased. In the nebulae, however, because of the extreme smallness of W, the number of these transitions was many times less than the number of transitions of the kind $1 \to 3 \to 2 \to 1$. Since, in envelopes of small radius, the dilution factor is not so small as in the nebulae, transitions of the kind $1 \to 2 \to 3 \to 1$ are made in them approximately as often as those of the kind $1 \to 3 \to 2 \to 1$.

When applied to the hydrogen atom, this means that, in envelopes of small radius, besides the processes of transformation of L_c quanta into quanta of lower frequency, the converse processes, i. e. the production of L_c quanta from quanta of lower frequency (for instance Lyman α + Balmer continuum $\to L_c$, etc.), occur also. The possibility of such processes is due to the fact that, because of the high degree of excitation of the atoms in the envelopes, their optical thicknesses, beyond the limits of the subordinate series, are comparable with unity, and the optical thicknesses in the lines of the subordinate series exceed unity.

It is clear that such a complication of the process of radiation of the envelopes, in comparison with the process of radiation of the nebulae, requires a considerable modification in ZANSTRA's method. For simplicity, let us consider the case where the optical thickness of the envelope beyond the limit of the Lyman series is much greater than unity. In this case we can suppose that, from every L_c quantum emitted by the star, there is formed in the envelope, sooner or later, a Balmer quantum, which leaves the envelope. It is evident that the ionisation of the atom from an excited state and the subsequent appearance of L_c quanta can only postpone the formation of the Balmer quantum, since all L_c quanta formed in the envelope will be absorbed in it again. However, an atom which is in the second state after emitting a Balmer quantum may be again transferred into the ionised state by the radiation of the star (immediately after emitting the Balmer quantum or after a number of scatterings of Lyman α quanta). It is easy to see that this leads to the emission of another Balmer quantum by the envelope. This process may be repeated several times, until finally a Lyman α

quantum, which escapes from the envelope, is formed by a transition of the atom from the second state to the first. Thus we reach the conclusion that Balmer quanta can be formed in the envelope not only from quanta in the Lyman continuum from the star, as in the nebulae, but also from those quanta in the Balmer continuum from the star which are absorbed by the envelope.

Taking these considerations into account (they refer, of course, not only to hydrogen atoms, but to other atoms also), we must write, instead of equation (29.1),

$$\left(1 + \frac{N_2}{N_1}\right) \int\limits_{x_0}^{\infty} \frac{x^2\,dx}{e^x - 1} = \sum \frac{x_i^3 A_i}{e^{x_i} - 1}\,, \tag{29.5}$$

where N_2/N_1 is the ratio of the number of quanta beyond the limit of the second series of the atom, emitted by the star and absorbed by the envelope, to the number emitted by the star beyond the limit of the principal series.

It is fairly difficult to estimate the ratio N_2/N_1. This was first done by Sh. G. Gordeladze on the basis of the theory of the excitation and ionisation of atoms in envelopes of small radius, developed by V. A. Ambartsumyan [6]. Sh. G. Gordeladze [47] has found that, when equation (29.1) is used instead of equation (29.5), temperatures considerably higher than the true values may be obtained, and the error is the larger, the higher the ionisation potential of the atom. However, V. A. Ambartsumyan's theory refers only to stationary envelopes or to those moving without a velocity gradient. If a velocity gradient is present in the envelope, the ratio N_2/N_1 is greatly decreased, and therefore the effect in question is less important.

2. **The continuous spectrum.** In characterising the radiation of Wolf-Rayet, P Cygni and Be-type stars, not only the temperatures found from the bright lines, but also the colour temperatures which represent the relative distribution of energy in the visible part of the star's spectrum, are of great importance. Since these stars are very distant from us, the galactic absorption of light has a great effect on the energy distribution in their spectra; as is well known, it causes a "reddening" of objects. The problem of taking into account this effect forms a serious difficulty in the determination of the colour temperatures of the stars.

The colour temperatures of Wolf-Rayet stars have been determined by B. A. Vorontsov-Vel'yaminov [172]. He found that, in the range 3800 to 4800 Å, the mean temperature of six Wolf-Rayet stars is 8400° in the scale where $T = 12{,}000°$ for stars of class A 0. After correcting for the absorption of light in space, the mean temperature was found

to be 13,000°. The temperatures of the individual stars (in thousands of degrees) are given in the last column of Table 25.

The spectrophotometric gradients of 39 Wolf-Rayet, O, and B-type stars have been determined by W. PETRIE [122]. To ascertain the effect of interstellar reddening, the gradients were compared with the intensities of interstellar absorption lines. After correcting the gradients for interstellar reddening, the following results were obtained for four groups of stars. The last two columns in Table 26 give the mean colour temperatures of the stars, temperatures of 18,000° and 15,000° respectively being taken for stars of type A 0.

Table 26

Group	Gradient	T_1	T_2
WR5 — WR8	—0·05	19,000°	16,000°
O5 — O6	—0·42	90,000	39,500
O7 — O9	—0·37	58,000	32,500
B0 — B2	—0·14	24,000	18,500

A spectrophotometric investigation of B-type stars has been carried out by D. BARBIER and D. CHALONGE [16]. They reached the conclusion that B-type stars with emission are, on the average, redder than B-type stars without emission.

Two important results follow from the work discussed above:

(1) Stars with bright lines in their spectra have lower colour temperatures than stars of the same spectral class without bright lines. This difference is small for stars of classes Be and B, but very large for stars of Wolf-Rayet type and those of class O.

(2) The colour temperatures of stars with bright spectral lines are lower than their temperatures as found from the bright lines. This discrepancy is small for Be stars, but very large for Wolf-Rayet stars.

The explanation of these observational results is based on the presence of envelopes in Wolf-Rayet and Be stars. The transformation of high-frequency radiation from the stars into quanta of lower frequencies, which takes place in these envelopes, leads to the emission by the envelope not only of energy in spectral lines, but also of energy in the continuous spectrum. Here the energy emitted by the envelope in the visible part of the continuous spectrum is comparable with the energy emitted by the star itself in that part of the spectrum. Since the radiation of the envelope corresponds to a lower temperature than does the radiation of the star, the colour temperatures of stars with envelopes will consequently be found to be lower than those of stars

without envelopes. On the other hand, the temperatures found from the bright lines are fairly close to the temperature of the star itself. This explains the discrepancy between the colour temperatures and Zanstra temperatures of stars with envelopes. It is clear that this discrepancy will be the greater, the thicker the envelope. This is why it is so marked in Wolf-Rayet stars, but less marked in Be stars.

Let us consider the more detailed interpretation of the continuous spectrum of a Be star [46]. As has been said above, the presence of envelopes in these stars does not prevent our seeing the absorption lines formed in the reversing layer of the star itself. This indicates the almost complete transparency of the envelopes of Be stars in the visible part of the spectrum. Since the absorption in the envelopes of Be stars is mainly due to hydrogen atoms, we may suppose that the optical thicknesses of these envelopes, beyond the limits of the subordinate series of the hydrogen atom, do not exceed unity (though they may be a considerable fraction of unity).

Thus the following processes occur in the envelopes of Be stars. Under the action of the radiation from the star, ionisation of hydrogen atoms takes place, both from the ground state and from excited states. The envelope thereby absorbs almost completely the radiation of the star beyond the limit of the Lyman series, and some of that beyond the limits of the other series. After the ionisation, there follow captures of electrons by protons and free-free transitions of electrons in the fields of protons. The quanta beyond the principal series limit, formed in these processes, are again almost completely absorbed in the envelope. However, the majority of the quanta beyond the limits of the subordinate series leave the envelope unhindered.

Let $n_e n_+ \varepsilon_{i\nu}' \, d\nu$ be the amount of energy emitted in the frequency interval from ν to $\nu + d\nu$ in 1 cm^3 and 1 second by captures of electrons by protons into the ith level. Next, let $n_e n_+ \varepsilon_\nu'' \, d\nu$ be the amount of energy emitted in the frequency interval from ν to $\nu + d\nu$ in 1 cm^3 and 1 second by free-free transitions of electrons in the fields of protons. Quantum mechanics gives the following expressions for the quantities $\varepsilon_{i\nu}'$ and ε_ν'':

$$\varepsilon_{i\nu}' = \frac{2^9 \pi^5}{(6\pi)^{3/2}} \left(\frac{m}{k T_e}\right)^{3/2} \frac{e^{10}}{m^2 c^3 h^2} \frac{1}{i^3} e^{(\chi_i - h\nu)/kT_e}, \tag{29.6}$$

$$\varepsilon_\nu'' = \frac{2^7 \pi^3}{(6\pi)^{3/2}} \frac{e^6}{m^2 c^3} \left(\frac{m}{k T_e}\right)^{1/2} e^{-h\nu/kT_e}, \tag{29.7}$$

where χ_i is the ionisation potential from the ith state, and T_e is the temperature of the electron gas. Formula (29.6) is valid for frequencies satisfying the condition $h\nu > \chi_i$. It is evident that the total amount

of energy emitted in 1 cm^3 and 1 second in the frequency interval from ν to $\nu + \mathrm{d}\nu$ is

$$n_e\, n_+\, \varepsilon_\nu\, \mathrm{d}\nu = n_e\, n_+ (\varepsilon_\nu'' + \sum_{i=j}^{\infty} \varepsilon_{i\nu}')\, \mathrm{d}\nu\,. \tag{29.8}$$

By means of formulae (29.6) and (29.7), we find for ε_ν

$$\varepsilon_\nu = \frac{2^7\,\pi^3}{(6\,\pi)^{3/2}} \left(\frac{m}{k\,T_e}\right)^{1/2} \frac{e^6}{c^3\, m^2} \left(1 + \frac{2\,h\,\nu_0}{k\,T_e} \sum_{i=j}^{\infty} \frac{1}{i^3}\, e^{\chi_i/kT_e}\right) e^{-h\nu/kT_e}\,. \tag{29.9}$$

Here we must take $j = 2$ in the frequency interval from the limit of the Balmer séries to that of the Lyman series, $j = 3$ in the frequency interval from the limit of the Paschen series to that of the Balmer series, etc.

Since the envelope is supposed transparent to its own radiation beyond the limits of the subordinate series, the amount of energy emitted by the entire envelope in the frequency ν is

$$E_{\nu,\,\mathrm{env}} = 4\,\pi\,\varepsilon_\nu \int_{r_0}^{\infty} n_e\, n_+\, r^2\, \mathrm{d}r\,, \tag{29.10}$$

where r_0 is the radius of the inner boundary of the envelope. In formula (29.10) it is assumed for simplicity that the envelope is spherically symmetrical and the electron temperature does not vary within it.

In order to calculate the integral (29.10), we must know the degree of ionisation of the atoms in the various parts of the envelope. We assume at first that the optical thickness of the envelope, beyond the limit of the Lyman series, is less than unity, i. e.

$$\tau_0 = k_{\nu_0} \int_{r_0}^{\infty} n_1\, \mathrm{d}r < 1\,. \tag{29.11}$$

In this case, we can use the ionisation formula derived for the case of the nebulae:

$$\frac{n_e\, n_+}{n_1} = W\, p \sqrt{\frac{T_e}{T_*}}\, \frac{(2\,\pi\, m\, k\, T_*)^{3/2}}{h^3}\, e^{-h\nu/kT_*}\,, \tag{29.12}$$

where $W = \frac{1}{4}\,(r_0/r)^2$ and p is the proportion of captures into the first level. Using formula (29.12), we find for τ_0

$$\tau_0 = \frac{k_{\nu_0}}{p} \sqrt{\frac{T_*}{T_e}}\, \frac{h^3}{(2\,\pi\, m\, k\, T_*)^{3/2}}\, e^{h\nu_0/kT_*}\, \frac{4}{r_0^2} \int_{r_0}^{\infty} n_e\, n_+\, r^2\, \mathrm{d}r\,. \tag{29.13}$$

We can now substitute the value of the integral $\int\limits_{r_0}^{\infty} n_e\, n_+\, r^2\, \mathrm{d}r$ from the relation (29.13) in formula (29.10). Using the familiar expressions for ε_ν and k_{ν_0}, we obtain

$$E_{\nu,\,\mathrm{env}} = 4\,\pi^2\, r_0^{\,2}\, \tau_0\, \frac{2\,h\,\nu_0^{\,3}}{c^2}\, e^{-h\nu_0/kT_*}\, \frac{kT_*}{h\nu_0}\, p \left(\frac{1}{2} + \frac{h\,\nu_0}{k\,T_e} \sum_{i=j}^{\infty} \frac{e^{\chi_i/kT_e}}{i^3}\right) e^{-h\,\nu/kT_e}\,. \tag{29.14}$$

Formula (29.14) has a simple physical significance. It is easy to see that the quantity

$$4\,\pi^2\, r_0^{\,2}\, \tau_0\, \frac{2\,h\,\nu_0^{\,3}}{c^2}\, e^{-h\nu_0/kT_*}\, \frac{k\,T_*}{h}$$

is that part of the energy from the star, beyond the limit of the Lyman series, which is absorbed by the envelope, and the quantity

$$p\, \frac{\mathrm{d}\nu}{\nu_0} \left(\frac{1}{2} + \frac{h\,\nu_0}{k\,T_e} \sum_{i=j}^{\infty} \frac{1}{i^3}\, e^{\chi_i/kT_e}\right) e^{-h\nu/kT_e}$$

is the ratio of the energy emitted by the envelope in the frequency interval from ν to $\nu + \mathrm{d}\nu$ to the total energy emitted by the envelope (to the same accuracy as that of the derivation of the ionisation formula). But the total energy emitted by the envelope must be equal to the part of the energy from the star in the Lyman continuum that is absorbed by the envelope. Hence, by multiplying these two quantities, we obtain $E_{\nu,\,\mathrm{env}}\, \mathrm{d}\nu$.

If the law of variation of density in the envelope is known, the quantities τ_0 and $E_{\nu,\,\mathrm{env}}$ can be expressed in terms of the density at the inner boundary of the envelope. Let us assume, for instance, that the density in the envelope decreases inversely as the square of the distance from the centre of the star. Then, taking into account the fact that the envelope is completely ionised for $\tau_0 < 1$ (i. e. the inequality $n_+/n_1 \gg 1$ holds throughout the envelope), we have

$$n_+ = n_e = n_e^{\,0}\, (r_0/r)^2\,. \tag{29.15}$$

Substituting n_e and n_+ from the relation (29.15) in formulae (29.13) and (29.10), we obtain

$$\tau_0 = \frac{k_{\nu_0}}{p} \sqrt{\frac{T_*}{T_e}}\, \frac{h^3}{(2\,\pi\, m\, k\, T_e)^{3/2}}\, e^{h\nu_0/kT_*}\, 4\,(n_e^{\,0})^2\, r_0 \tag{29.16}$$

and

$$E_{\nu,\,\mathrm{env}} = 4\,\pi\,\varepsilon_\nu\, (n_e^{\,0})^2\, r_0^{\,3}\,. \tag{29.17}$$

The case where $\tau_0 > 1$ is more difficult to examine. In this case, the degree of ionisation of the atoms in the envelope is determined from the theory of radiative transfer in the envelope. However, this theory is more complex than that of the transfer of L_c radiation in the nebulae, since, in envelopes of small radius with $\tau_0 > 1$, ionisation from excited states is important. We shall not pause to consider this theory here. We shall merely emphasise the fact that, in the case mentioned, the energy emitted by the envelope may be considerably more than that given by formula (29.14) with $\tau_0 = 1$.

According to what was said earlier, the spectrum of a Be star can be regarded as the result of the superposition of the spectrum of the envelope on that of a B star; thus the energy emitted by a Be star in the frequency ν can be calculated from the formula

$$E_\nu = E_{\nu,*} + E_{\nu,\mathrm{env}}\,, \tag{29.18}$$

where $E_{\nu,*}$ is the energy emitted by a B star in the frequency ν. We shall suppose that a B star emits like a black body of temperature T_*, i. e.

$$E_{\nu,*} = 4\,\pi^2\,r_0^{\,2}\,\frac{2h\nu^3}{c^2}\,\frac{1}{e^{h\nu/kT_*}-1}\,. \tag{29.19}$$

Comparing formulae (29.14) and (29.19), we find that, for a temperature of the star of the order 20,000° to 30,000°, and for values of τ_0 not much less than unity, the ratio $E_{\nu,\mathrm{env}}/E_{\nu,*}$ in the visible part of the spectrum will be of the order of 0·5 (smaller in the violet part of the spectrum and greater in the red). As has been said already, for $\tau_0 > 1$ this ratio may become quite considerable. Consequently, the energy distribution in the spectrum of a Be-type star may differ quite markedly from that in the spectrum of a B-type star.

From the formulae given above we can confirm that, as the intensity of ejection of matter from the star increases, the following changes should be observed in its radiation.

First of all, the apparent brightness of the star should increase. This is seen from formulae (29.17) and (29.18). Applying these formulae to the radiation of wavelength $\lambda = 5550$ Å, to which our eyes are most sensitive, we can approximately calculate the increase in the brightness of the star as $n_e^{\,0}$ increases.

Next, the colour temperature of the star should decrease. As we know, the colour temperature T_c is determined from the condition

$$\frac{\mathrm{d}}{\mathrm{d}\nu}\log_e E_\nu = \frac{3}{\nu} - \frac{h}{k\,T_c}\,\frac{1}{1-e^{-h\nu/kT_c}}\,. \tag{29.20}$$

where the value of $\mathrm{d}(\log_e E_\nu)/\mathrm{d}\nu$ is found from observation. If we suppose that the radiation of the envelope in the visible part of the spectrum is considerably stronger than that of the star, then, substituting $E_{\nu,\mathrm{env}}$ for E_ν and taking into account the fact that

$$E_{\nu,\,\mathrm{env}} \sim e^{-h\nu/kT_e}\,,$$

we find for the determination of T_c

$$\frac{h\,\nu}{k}\left(\frac{1}{T_c}-\frac{1}{T_e}\right)=3 \tag{29.21}$$

(we have neglected the term $e^{-h\nu/kT_c}$ in comparison with unity). It is seen from formula (29.21) that the colour temperature of the envelope is much lower than its electron temperature. For example, for $T_e = 20{,}000°$ we have $T_c = 6700°$ in the region of the H_β line. This means that, as the part played by the radiation of the envelope increases, the colour temperature of the star will decrease from the value $T_c = T_*$ to the value given by formula (29.21).

Finally, as the outflow of matter from the star becomes more intense, the intensity discontinuity at the limit of the Balmer series should decrease. We have supposed above that the energy distribution in the spectrum of the star is given by Planck's formula. In reality this is not so. As we know from the theory of photospheres, the energy distribution in the spectrum of a B-type star differs very markedly from the Planckian distribution, its characteristic feature being the intensity discontinuity at the limit of the Balmer series. Since the absorption coefficient of the hydrogen atom beyond the limit of the Balmer series (i. e. for $\nu > \nu_2$) is greater than it is up to that limit (i. e. for $\nu < \nu_2$), the intensity of radiation in the star's spectrum beyond the limit of the Balmer series is less than it is up to that limit. The quantity

$$D = \log_{10}\left(E_{\nu<\nu_2}/E_{\nu>\nu_2}\right) \tag{29.22}$$

is usually found from observation; it is positive for "normal" B-type stars, in agreement with what we have just said. However, the radiation of the envelopes has the opposite property; for them, the quantity D_{env}, determined by the formula

$$D_{\mathrm{env}} = \log_{10}\left[\left(1+2\,\frac{h\,\nu_0}{kT_e}\sum_{i=3}^{\infty}\frac{e^{\chi_i/kT_e}}{i^3}\right)\Big/\left(1+2\,\frac{h\,\nu_0}{kT_e}\sum_{i=2}^{\infty}\frac{e^{\chi_i/kT_e}}{i^3}\right)\right], \tag{29.23}$$

is negative. Hence the superposition of the radiation of the envelope on that of the star leads to a decrease in the intensity discontinuity at the limit of the Balmer series. As the radiation of the envelope becomes more intense, the value of D for a Be star may even change from positive to negative.

It must also be remarked that, as the energy radiated by the envelope in the continuous spectrum increases, the energy radiated by it in the spectral lines must increase also. If the same conditions existed in the envelopes of Be stars as in the nebulae, then, for a given electron temperature, the ratio of the intensity of any line to that at any point in the continuum (excepting the lines and continuum of the principal series) would always remain constant. However, envelopes of small radius, unlike the nebulae, are opaque to radiation in the lines of the subordinate series, and ionisation from excited states may occur in them. This leads to a redistribution of energy between the lines and the continuum, whereby, as is easily seen, the radiation in the continuum is strengthened in comparison with that in the lines. The corresponding calculations can easily be performed by using the theory given in Section 28.4, and the formulae in the present section.

The theoretical conclusions just given concerning the properties of the spectra of Be stars are in good agreement with observation. This is most obvious from the work of V. G. Gorbatskiĭ on one of the most remarkable of the Be stars, γ Cassiopeiae [45]. This bright star (of the second magnitude) exhibits irregular variations of brightness and spectrum. Particularly violent changes occurred in the period from 1936 to 1941. The study of extensive observational data concerning this star has led to the conclusion that, in the period mentioned, the intensity of outflow of matter from the star on three occasions increased and then decreased again (roughly speaking, the star ejected three envelopes in succession). Each time the outflow of matter from the star was intensified, its brightness increased, its colour temperature decreased, the intensity discontinuity D at the limit of the Balmer series diminished, and the intensity of the Balmer lines increased (Fig. 73, where H_δ is shown). A detailed comparison of theory and observation enabled V. G. Gorbatskiĭ to determine a number of characteristics of γ Cassiopeiae. According to his calculations, the radius of the photosphere is equal to three solar radii, the temperature of

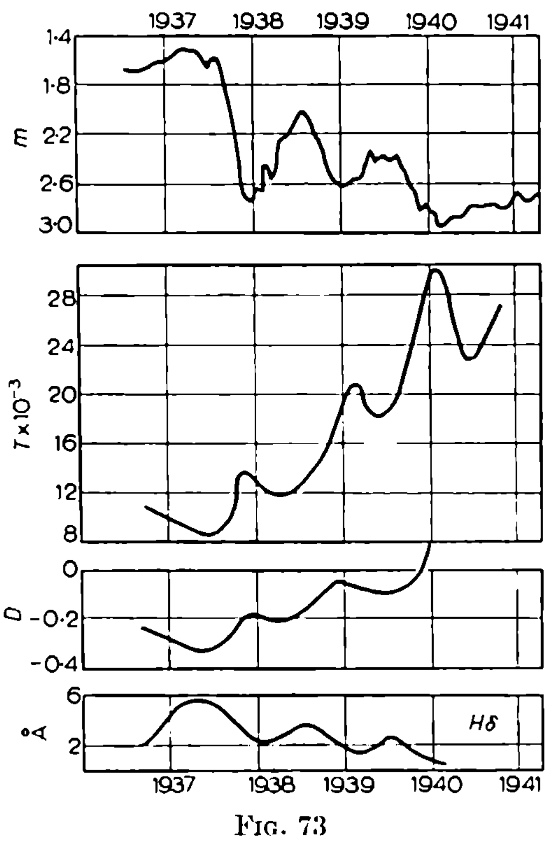

Fig. 73

the star is 34,000°, the number of free electrons in 1 cm³ at the star's surface is 10^{12} to 10^{13}, and the mean electron temperature of the envelope is 15,000° to 20,000°.

Let us now turn to Wolf-Rayet type stars. The radiation of the envelopes of these stars is produced in fundamentally the same way as that of the envelopes of Be-type stars. However, Wolf-Rayet stars are much hotter than Be stars. For this reason, not only hydrogen atoms but also those with higher ionisation potentials take part in the formation of the continuous spectrum of Wolf-Rayet stars.

Another important characteristic of the process of radiation by the envelopes of Wolf-Rayet stars is the large part played by the scattering of radiation by free electrons. In order to demonstrate this, let us calculate the optical thickness of the "atmosphere" of a Wolf-Rayet star due to electron scattering (we recall that the envelope of a Wolf-Rayet star can be artificially divided into a "photosphere" and an "atmosphere"). Denoting this thickness by τ_e, we have

$$\tau_e = \sigma_0 \int_{r_0}^{\infty} n_e \, dr \,, \qquad (29.24)$$

where σ_0 is the scattering coefficient referred to one electron, and r_0 is the radius of the photosphere. Let us assume, as before, that the density of matter in the envelope of a Wolf-Rayet star diminishes in inverse proportion to the square of the distance from the centre of the star. Then we find, instead of (29.24),

$$\tau_e = \sigma_0 \, n_e^0 \, r_0 \,. \qquad (29.25)$$

To estimate τ_e, we take $r_0 = 5\, r_\odot$, $n_e^0 = 1{\cdot}5 \times 10^{12}$. It is known that $\sigma_0 = 0{\cdot}67 \times 10^{-24}$. The formula (29.25) gives $\tau_e = \frac{1}{3}$. We see that the optical thickness of the "atmosphere" of a Wolf-Rayet star due to electron scattering is very large. The lower limit of the "atmospheres" of these stars is apparently determined, in fact, by the condition that τ_e is of the order of unity there.

For comparison, we may calculate the optical thickness of the "atmosphere" beyond the limit of the Lyman series. We use formula (29.16) for this purpose. Taking $r_0 = 5\, r_\odot$, $n_e^0 = 1{\cdot}5 \times 10^{12}$, $T_* = 50{,}000°$ and using the fact that $k_{\nu_0} = 0{\cdot}5 \times 10^{-17}$, we obtain $\tau_0 = 0{\cdot}03$. Consequently $\tau_e \gg \tau_0$, i. e. in the atmospheres of Wolf-Rayet stars the scattering of radiation by free electrons plays a greater part than the absorption by hydrogen atoms. As we penetrate into the photosphere, the electron scattering gradually becomes less important in comparison with atomic absorption. This is due to the fact that the volume coefficient of scattering by free electrons is proportional to n_e, whilst the volume coefficient of absorption by the atoms is proportional to $n_e \, n_+$.

Because of the large part played by electron scattering in the envelopes of Wolf-Rayet stars, the radiation of these stars has a number of interesting properties. The most important is the polarisation of the star's radiation. Of course, this property is found not only in Wolf-Rayet stars, but also in other stars with a fairly high temperature and a fairly low density in the outer layers. To ascertain the degree of polarisation of stellar radiation, the problem has been solved of the transfer of radiation in a pure electron photosphere composed of plane-parallel layers [30, 156]. It has been found that the degree of polarisation of the radiation from a star having such a photosphere increases from zero at the centre to 12·5 % at the limb, the radiation being polarised along the radius. It is evident that the radiation coming from the whole disc of a spherically symmetrical star will be unpolarised. Hence, to reveal the effect in question, it is necessary to observe eclipsing variables, one of whose components is a hot star and the other a cool star. In this case, when the hot star is obscured by its cool companion, the radiation of the system will be polarised to some extent, though only very slightly. This effect, predicted by the theory, was actually discovered in subsequent observations on a number of eclipsing variables.

It is particularly interesting that in these observations a completely new phenomenon was discovered also, namely the polarisation of the light of the stars outside eclipse and even the polarisation of the light of single stars [38, 39, 56, 60]. The greatest degree of polarisation (of the order of 6 %) is shown by some Wolf-Rayet stars. The explanation of this phenomenon has not yet been found. It is possible that it is partly due to the scattering of light by free electrons in envelopes not having spherical symmetry. It has also been suggested that the light of the stars becomes polarised in passing through the interstellar medium.

3. **Stars of late types with bright lines.** Besides Wolf-Rayet, P Cygni and Be-type stars, which belong to the early spectral classes, emission lines are also found in the spectra of stars of late types. Among these are, first of all, the long-period variables, the majority of which show in their spectra bright lines of hydrogen, ionised iron (near maximum light) and neutral iron (near minimum light). Other such stars are those of the Z Andromedae type. Their spectra consist of a combination of late-type continuous and absorption spectra and a bright-line spectrum belonging to atoms with very high ionisation potentials (for example, He II).

The Balmer decrement in the spectra of the long-period variables is very interesting. The appropriate data are given in Table 27. We see that the Balmer decrement in the spectra of Me-type stars is very anomalous. However, this is not due to the operation of some mechanism, unknown to us, for the excitation of atoms. As G. A. Shaĭn [137]

Table 27

Type	Me	Se	Ne
H_α	2	15	10
H_β	2	12	10
H_γ	20	5	5
H_δ	30	3	2

has shown, the cause of the anomaly in the Balmer decrement is the absorbing action of titanium oxide. In the spectra of Se and Ne-type stars, the titanium oxide bands are absent, and the Balmer decrement is quite normal. It is also of interest that, in the spectra of the long-period variables, the bright hydrogen lines are divided into a number of components. It has recently been shown that this unusual structure of the bright lines is again explained by the screening of the hydrogen radiation in the atmospheres of the stars, due this time to atoms of un-ionised metals. From these facts, the conclusion can be drawn that the emission lines are formed mainly in deeper layers of the atmosphere than are the absorption lines.

Very valuable information can also be obtained by considering the radial velocities of long-period variables. As is well known, a variation takes place in the displacements of spectral lines in these stars, besides the oscillation in their brightness. It is interesting that the displacements of the emission and absorption lines are not the same. This again indicates that the emission and absorption occur in different layers of the star's atmosphere. A statistical discussion of the radial velocities of long-period variables leads to the following results:

(1) the difference of the radial velocities found from bright and from dark lines is always negative ($v_e - v_a < 0$);

(2) the K-term determined from the bright lines is about —15 km/sec;

(3) the K_a-term determined from the dark lines is about zero.

It follows from these data that the layer in which the bright lines are formed is moving in the direction of the observer. In other words, an outflow of matter takes place from long-period variables. This conclusion was first reached by G. A. Shaĭn [140]. Some doubts are occasioned here by the fact that $K_a \approx 0$. However, for a number of reasons, this value of K_a seems to be too large. If this is so, the hypothesis that matter is ejected from long-period variables is fairly plausible. Here the process of ejection of matter must be not steady but variable. The ejected matter itself must undergo a considerable retardation.

The results mentioned are very important for an understanding of the nature of late-type stars with bright lines. In particular, they are very useful in resolving the following important question which occurs in connection with these stars: what is the reason for the appearance of bright lines in the spectra of such cool stars? The ultra-violet energy from a star whose temperature is of the order of 2000° to 3000° is clearly insufficient for bright lines to be formed from this energy.

In order to solve this problem, we rephrase it as follows: let there be a hot star surrounded by an extended and rarefied envelope, in which bright lines are formed as a result of the transformation of ultra-violet energy from the star. The question is whether the envelope can give under these conditions, besides the bright lines, a continuous and line spectrum of late type.

We have previously shown that, as the optical thickness of the envelope increases, the colour temperature of the object decreases. This explains the discrepancies between the Zanstra and colour temperatures of Be and Wolf-Rayet type stars. In essentials we are concerned with similar discrepancies, although these are even more strongly marked, in the case of late-type stars with bright lines. It may therefore be supposed that late-type stars with bright lines differ from early-type stars with bright lines by the even greater optical thickness of their envelopes. In other words, late-type stars with bright lines should apparently be regarded as hot stars surrounded by very extended envelopes of great optical thickness in the continuous spectrum. The processes occurring in such envelopes can be briefly described as follows. The inner parts of the envelope completely absorb the radiation of the star and transform it into quanta of low frequency. A late-type continuous spectrum is thereby formed, and bright lines of various atoms also appear. The outer parts of the envelope are scarcely reached by the radiation of the star, and they come under the action of the low-temperature radiation of the envelope itself. In these parts there are un-ionised metal atoms and molecular compounds. The late-type absorption spectrum is formed here.

It is correct to say that the hypothesis given above is confirmed in its general outlines by the observational data. For example, the fact that the bright lines observed in the spectra of long-period variables are formed in deeper layers of the envelope than the absorption lines and bands is in complete agreement with this hypothesis. We recall that this is shown by the fact that the bright lines are screened by molecular bands and the lines of un-ionised metals. Further, it must be remarked that, at the time near maximum brightness, i.e. when this screening is least, the emission spectra of long-period variables are very similar to those of Be-type stars, and novae at the time when

the bright lines appear. This fact may be regarded as a direct indication of a common mechanism of the excitation of emission lines in all the cases mentioned.

The changes in the spectra of Z Andromedae type stars are also in favour of the hypothesis mentioned. These changes are so marked that a late-type star is converted into an early-type one. This happened to the star Z Andromedae itself in 1939. These conversions can be explained by a great change in the force of ejection of matter from the star, and a consequent change in the thickness of the envelope. The spectra of these stars are sometimes a combination of two continuous spectra, one early and the other late. In such cases, according to the proposed interpretation, the spectra of the star and of the envelope are comparable in brightness. Here, roughly speaking, the red end of the combined spectrum belongs to the envelope, and the violet end to the star.

According to the hypothesis in question, late-type stars with emission lines may be converted into late-type stars without emission lines by a further increase in the optical thickness of the envelope. The "typical" cool supergiants are apparently formed in this way. In this connection it must be pointed out that the masses and luminosities of M-type supergiants are the same as those of O and B-type stars. This compels us to suppose that there is no essential difference in the internal structure of these stars.

For the sake of completeness, we may say that other hypotheses also have been advanced to explain the origin of late-type spectra with emission lines. The hypothesis deserving most attention is that of L. BERMAN, who regards Z Andromedae type stars as double stars consisting of a blue and a red component. However, no convincing proof has yet been obtained that any of these stars is actually double.

4. Stellar associations. All the stars with bright spectral lines discussed above have very high luminosities. However, besides these, dwarf stars with bright lines in their spectra have also been revealed by observation. The most remarkable of these are the T Tauri type variables. These stars mainly belong to the spectral classes G to M and have bright lines of H, Ca II, Fe II, etc. In some cases, absorption lines are seen on the violet side of the bright lines. This compels us to suppose that an ejection of matter is taking place from T Tauri type stars, as it is from other stars with similar line contours.

The spatial distribution of T Tauri type stars is very peculiar. The consideration of this distribution was the starting point for V. A. AMBARTSUMYAN's discovery and study of a new kind of stellar system, the stellar associations [10].

V. A. Ambartsumyan pointed out that almost all T Tauri type stars known (up to 1947) are found in two very small regions of the sky. In one of these regions, in Taurus and Auriga, there are concentrated eight T Tauri type stars and about 40 other dwarf stars with emission lines in their spectra. As was subsequently shown by P. N. Kholopov, the majority of the latter stars are also T Tauri type variables. The distance of this group of stars is of the order of 100 parsecs, and the diameter of the group is of the order of 25 parsecs. Another group of T Tauri type variables, numbering seven objects, is located in the constellations of Aquila and Ophiuchus; it is apparently further from us than the first group.

V. A. Ambartsumyan next noticed the fact that hot stars (of types O and B) also have a tendency to congregate. He and his collaborators have discovered more than 20 groups of such stars. It is particularly interesting that the majority of the Wolf-Rayet and P Cygni type stars, and many Be stars also, appeared in these groups. A characteristic example of a group of hot stars is the system surrounding the double cluster χ and h Persei. This system includes some tens of B-type supergiants, among them five P Cygni type stars. The cluster is the nucleus of the system. The diameter of the system is of the order of 170 parsecs, and its distance is about 2000 parsecs.

These groups of stars were called by V. A. Ambartsumyan **stellar associations** (the T and O associations respectively). It must be emphasised that there is a great difference between associations and clusters. Whereas the density in a cluster considerably exceeds the mean density of the star field surrounding it, the density of an association is small compared with the field density. However, associations are clearly distinguished by their *partial density*, i. e. the density due to stars of a definite class (dwarf stars with bright spectral lines in one case, O and B-type supergiants in the other). It is for this reason that the associations were discovered as a result of observation.

The conclusions drawn by V. A. Ambartsumyan from the very fact of the existence of stellar associations are extremely important. It is utterly improbable that the associations were formed by chance encounters of stars. Consequently, we must suppose that the stars forming an association have been related during their entire existence. However, calculation shows that this relation is very insecure, and the associations should rapidly disperse owing to the tidal action of the general gravitational field of the Galaxy. Nevertheless, they have not yet dispersed. This evidently indicates that the associations, and therefore the stars in them, were formed comparatively recently. According to estimates which have been made, their age does not exceed 10^7 years, whereas the mean age of the stars is known to be 10^9 to 10^{10} years. Hence

the stellar associations must be acknowledged to be very young formations.

This conclusion, that the stellar associations are extremely young, which was derived from dynamical considerations, is confirmed by astrophysical data. As we have already said, there are in the associations Wolf-Rayet, P Cygni, Be and other types of stars which eject matter. The amount of matter ejected by each such star per year is 10^{-6} to 10^{-5} times the Sun's mass. It is clear that this process cannot continue for more than 10^6 to 10^7 years. Consequently the stars which eject matter must pass very rapidly from one stable state with a large mass to another stable state with a smaller mass. However, we know that stars with masses exceeding those of Wolf-Rayet, P Cygni and Be-type stars are exceedingly rare. Hence we must again conclude that the stars in question, and therefore the associations containing them, were formed directly from the pre-stellar phase of matter, and that this occurred quite recently.

Thus the formation of stars in our Galaxy, which began several thousand million years ago, is still continuing. This conclusion is especially important in that it contradicts theories of the simultaneous formation of all the stars.

The variable stars found in T associations show very irregular variations in brightness. It is of interest that the rapid increase which sometimes occurs in the brightness of these stars is often accompanied by the appearance of a peculiar "continuous emission", which is superposed on the normal absorption-line spectrum. As a result, the absorption lines are partly or wholly masked. Thus the increase in brightness often occurs because of the sudden addition to the normal thermal radiation of a non-thermal radiation having a continuous spectrum. It seems that in these cases powerful processes of energy release are occurring in the outer layers of these stars, and this results in an emission of non-thermal radiation.

All this indicates the existence of very violent processes in T Tauri type stars. It is natural to suppose that they also are stars "in course of formation".

It can hardly be doubted that the formation of stars in an earlier period in the life of the Galaxy took place through associations. Thus the process of the evolution of the Galaxy must be imagined as the gradual formation of groups of stars (i. e. associations) from the pre-stellar material. As has been said, the associations are dynamically unstable and rapidly disperse. As a result, the stars in the associations are intermixed with those formed previously. The process of evolution of the stars goes on at the same time. At the beginning of their existence the stars are unstable, like the associations, as is shown by the ejection

of matter from them. However, as time goes on, by ejecting matter, they acquire stability and pass into the class of "ordinary" stars. It is important in this connection that the stars when formed may be either giants or dwarfs (in consequence of the fact that they may be formed either through O associations or through T associations). In other words, the stars "enter" the spectrum-luminosity diagram not only at the head of the main sequence, but also in other parts of it, i. e. all along the main sequence.

Of the other results obtained in the study of stellar associations, the most important is that many double and multiple stars are found in the associations. Thus, for example, the majority of Wolf-Rayet type stars seem to be components of close binaries. Many of the T Tauri type stars have visual companions. These facts throw fresh light on the problem of the origin of double stars. Earlier, as is well known, it was supposed that double stars are formed either by the fission of single stars or by the capture of one star by another. However, we see now that even in the associations, i. e. among the youngest stars, there are many double and multiple systems. Consequently, it is incorrect to suppose that stars acquire companions in the course of their long lives. In reality, they are formed either single or multiple. It is possible that there is no fundamental difference between multiple stars and the solar system. In that case, the conclusion mentioned is of importance for the cosmogony of the solar system also.

It is also interesting that there exists a great difference between multiple systems in associations and those outside. Multiple systems outside associations usually have distances between the components which are of different orders of magnitude. For example, triple systems are constructed in such a way that two of the stars (A and B) are close to each other, whilst the third star (C) lies at a distance from them. The motion of the stars in this system can be approximately regarded as a Keplerian motion of the star B round the star A, and a Keplerian motion of the star C round the pair AB. Quadruple systems are similarly constructed. They usually consist of two pairs of close stars with a considerable distance between the pairs. It is found that only such systems are stable. If a multiple system were formed with the distances between the components all of the same order, it would rapidly disintegrate or be converted into a system with the distances between the components of different orders of magnitude (i. e. a system of the kind mentioned above). In the associations, however, multiple systems are found with the distances between the components all of the same order. V. A. Ambartsumyan and B. E. Markaryan call these Trapezium-type systems (in analogy with the well-known Trapezium in Orion, which, incidentally, belongs to the Orion association). It is clear that

the presence of Trapezium-type multiple systems in stellar associations is another proof that the associations are young formations [11].

The further investigation of stellar associations will undoubtedly lead to new results which are of great importance in cosmogony. Even in the next few years, it seems, we may expect the problem to be resolved of the nature of the matter from which the stars are formed.

PART VII.
THE INTERNAL STRUCTURE OF STARS

Chapter 30. Basic facts

1. **Statement of the problem.** The problem of the internal structure of stars is one of the most difficult problems of theoretical astrophysics. Many aspects of it are still unresolved. Many details have not been investigated at all. Hence we shall restrict ourselves to an exposition only of some basic facts necessary to gain an idea of the methods of solution of this problem. We shall pay no attention to the variations of particular models of the internal structure of individual types of star, which have been worked out on various more or less arbitrary assumptions. It seems to us that the value of a great number of schematic models developed in this way is extremely limited.

Let us consider steady (non-variable) stars in a state of equilibrium, regarding a star as spherically symmetrical, i.e. neglecting rotation, which causes a deviation from spherical symmetry. Then all the quantities characterising the physical state of the star (the temperature T, the density ϱ, the pressure p, the acceleration g due to gravity, the flux H of radiation, etc.) will be functions only of the distance r from the centre of the star. On this assumption, the equation of hydrostatic equilibrium is

$$\mathrm{d}p/\mathrm{d}r = -g\,\varrho\,, \tag{30.1}$$

where the acceleration due to gravity is

$$g = G\mathsf{M}(r)/r^2\,, \tag{30.2}$$

G being the constant of gravitation and $\mathsf{M}(r)$ the mass inside a sphere of radius r, and the total pressure $p = p_G + p_R$, where p_G is the gas pressure and p_R the radiation pressure. Next, it follows from the equation of transfer, on multiplying by $\cos\theta$ and integrating, that

$$\mathrm{d}p_R/\mathrm{d}r = -k\,\varrho\,H/c\,, \tag{30.3}$$

if we introduce the absorption coefficient k referred to unit mass (c is the velocity of light).

Let $\varepsilon(r)$ be the amount of energy evolved per unit mass in unit time by the sources of energy at a distance r from the centre of the star. Then the total energy released by the sources inside the sphere of radius r is

$$L(r) = \int_0^r 4\,\pi\,r^2\,\varepsilon\,\varrho\,\mathrm{d}r\,. \tag{30.4}$$

If the principal means of energy transfer inside the star is *by radiation* (as is true for the high temperatures existing inside stars), then

$$L(r) = 4\,\pi\,r^2\,H\,. \tag{30.5}$$

Thus equation (30.3) can be written in the form

$$\mathrm{d}p_R/\mathrm{d}r = -\,k\,\varrho\,L(r)/4\,\pi\,c\,r^2\,, \tag{30.6}$$

and equation (30.1), by (30.2), takes the form

$$\frac{\mathrm{d}(p_G + p_R)}{\mathrm{d}r} = -\,\frac{G\,\mathsf{M}(r)}{r^2}\,\varrho\,. \tag{30.7}$$

To these two basic equations (30.6) and (30.7), we must add equation (30.4), which we write in the form

$$\mathrm{d}L(r)/\mathrm{d}r = 4\,\pi\,r^2\,\varrho\,\varepsilon\,, \tag{30.8}$$

and the following obvious condition:

$$\mathrm{d}\mathsf{M}(r)/\mathrm{d}r = 4\,\pi\,r^2\,\varrho\,. \tag{30.9}$$

Let us now consider a gaseous star; the equation of state of the matter inside such a star is given by Boyle's Law,

$$p_G = R\,\varrho\,T/\mu\,, \tag{30.10}$$

where R is the gas constant and μ the molecular weight. The radiation pressure is given by the formula

$$p_R = \tfrac{1}{3}\,a\,T^4\,, \tag{30.11}$$

which holds for the internal layers of a star which are in a state of local thermodynamic equilibrium. Thus, for a gaseous star, the system of fundamental equations describing its equilibrium state takes the form

$$\left.\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{R}{\mu}\,\varrho\,T + \tfrac{1}{3}\,a\,T^4\right) = -\,\frac{G\,\mathsf{M}(r)}{r^2}\,\varrho\,,\\
&\frac{\mathrm{d}}{\mathrm{d}r}\left(\tfrac{1}{3}\,a\,T^4\right) = -\,\frac{k\,L(r)}{4\,\pi\,c\,r^2}\,\varrho\,,\\
&\mathrm{d}\mathsf{M}(r)/\mathrm{d}r = 4\,\pi\,r^2\,\varrho\,,\\
&\mathrm{d}L(r)/\mathrm{d}r = 4\,\pi\,r^2\,\varrho\,\varepsilon\,.
\end{aligned}\right\} \tag{30.12}$$

The solutions of this system of four equations for the four variables ϱ, T, $\mathsf{M}(r)$ and $L(r)$ must satisfy the following boundary conditions:

$$\left.\begin{array}{lllll} \text{for } r=0, & \varrho=\varrho_c, & T=T_c, & \mathsf{M}(r)=0, & L(r)=0; \\ \text{for } r=r_1, & \varrho=0, & T=0, & \mathsf{M}(r)=\mathsf{M}, & L(r)=L, \end{array}\right\} \tag{30.13}$$

where r_1 is the radius, M the mass and L the luminosity of the star, and the suffix c denotes the value of a quantity at the centre of the star.

We introduce dimensionless variables by expressing ϱ, T, k and ε in terms of their central values and $L(r)$ and $\mathsf{M}(r)$ in terms of the luminosity and mass of the whole star:

$$\left.\begin{array}{lll} T=T_c\,\tau, & \mathsf{M}(r)=\mathsf{M}\,\psi, & L(r)=L\,\eta, \\ \varrho=\varrho_c\,\sigma, & \varepsilon=\varepsilon_c\,\varepsilon', & k=k_c\,\varkappa, \end{array}\right\} \tag{30.14}$$

and put

$$r=r_1\,\alpha\,\xi, \tag{30.15}$$

where α is a parameter which is as yet arbitrary. It is then not difficult to show that the equations (30.12) reduce to the system

$$\left.\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\xi}\left(\sigma\tau+\frac{1-\beta_c}{\beta_c}\tau^4\right) &= -\frac{\sigma\psi}{\xi^2}, \\ \mathrm{d}\tau^4/\mathrm{d}\xi &= -\lambda_1\,\varkappa\,\sigma\,\eta/\xi^2, \\ \mathrm{d}\psi/\mathrm{d}\xi &= \lambda_2\,\sigma\,\xi^2, \\ \mathrm{d}\eta/\mathrm{d}\xi &= \lambda_3\,\sigma\,\varepsilon'\,\xi^2, \end{aligned}\right\} \tag{30.16}$$

where the dimensionless parameters are

$$\left.\begin{aligned} &\alpha=\frac{G\mathsf{M}}{(R/\mu)\,r_1\,T_c}, \quad \lambda_1=\frac{k_c\,L}{4\pi cG\mathsf{M}}\,\frac{\beta_c}{1-\beta_c}, \\ &\lambda_2=\frac{4\pi G^3\mathsf{M}^2}{(R/\mu)^4}\,\frac{\beta_c}{1-\beta_c}\,\tfrac{1}{3}\,a, \quad \lambda_3=\mathsf{M}\,\lambda_2\,\varepsilon_c/L, \end{aligned}\right\} \tag{30.17}$$

and $(1-\beta_c)/\beta_c=(p_R/p_G)_c$. The equations (30.16) have to be solved with the boundary conditions (30.13) or

$$\left.\begin{array}{lllll} \text{for } \xi=0, & \sigma=1, & \tau=1, & \psi=0, & \eta=0; \\ \text{for } \xi=1/\alpha, & \sigma=0, & \tau=0, & \psi=1, & \eta=1. \end{array}\right\} \tag{30.18}$$

The solution of the system (30.16) contains four parameters: λ_1, λ_2, λ_3 and $(1-\beta_c)/\beta_c$, and four arbitrary constants which are determined, for example, by the conditions (30.18) at the centre, which enable us to ex-

press these four constants in terms of λ_1, λ_2, λ_3 and $(1-\beta_c)/\beta_c$. Thus the solution satisfying the conditions (30.18) at the centre has the form

$$\sigma,\ \tau,\ \psi,\ \eta = f_i(\xi, (1-\beta_c)/\beta_c,\ \lambda_1,\ \lambda_2,\ \lambda_3) \quad (i = 1, 2, 3, 4). \qquad (30.19)$$

This solution, however, must also satisfy the conditions (30.18) for $\xi = 1/\alpha$, and this leads to four conditions of the form

$$\Phi_i(1/\alpha,\quad (1-\beta_c)/\beta_c,\quad \lambda_1, \lambda_2, \lambda_3) = 0 \quad (i = 1, 2, 3, 4)\ . \qquad (30.20)$$

Solving these four equations for the four dimensionless parameters λ_1, λ_2, λ_3 and $1/\alpha$ (the eigenvalues), we find expressions for them in terms of the ratio $(1-\beta_c)/\beta_c$. By equations (30.17), these expressions have the form

$$F_i(L,\ \mathsf{M},\ r_1,\ T_c,\ \varrho_c,\ k_c,\ \varepsilon_c,\ \mu) = 0 \quad (i = 1, 2, 3, 4)\ , \qquad (30.21)$$

since $(1-\beta_c)/\beta_c = \frac{1}{3}\, a\, T_c^3/(R/\mu)\varrho_c$. Determining ϱ_c and T_c from two of the equations (30.21), we find

$$\varrho_c,\ T_c = \Psi_i(L,\ \mathsf{M},\ r_1,\ k_c,\ \varepsilon_c,\ \mu) \quad (i = 1, 2)\ , \qquad (30.22)$$

i. e. the conditions determining the temperature and density at the centre of the star. Substituting these expressions (30.22) in the remaining two equations (30.21), we obtain two relations of the form

$$F_i(L,\ \mathsf{M}, r_1, k_c, \varepsilon_c, \mu) = 0 \quad (i = 1, 2)\ . \qquad (30.23)$$

The fundamental conclusions of the theory of the internal structure of stars follow from a discussion of equations (30.22) and (30.23).

2. **The basic empirical relations.** The observer characterises a star by three physical parameters: its mass M, its luminosity L and its radius r_1, each of which can be determined experimentally to a known degree of accuracy. The observations show that, for stars of the main sequence, giants and supergiants (see the spectrum-luminosity diagram, Fig. 46), there is an almost rigorous functional relation between the mass and the luminosity, of the form $L \sim \mathsf{M}^n$. According to the investigations of P. P. Parenago and A. G. Masevich [120], the form of this dependence is different for the various groups of stars mentioned: for the giants and supergiants, $n \approx 10/3$; for stars of the spectral classes O to G 4 (first part of the main sequence), $n = 3{\cdot}9$; for stars of classes G 7 to M (second part of the main sequence), $n \approx 2{\cdot}3$. For the white dwarfs, the subdwarfs and the subgiants, no reliable relation between the mass and the luminosity has been discovered.

Observation also reveals a correlation between the radius and the luminosity (or the mass), or between the luminosity and the effective temperature, since the luminosity is

$$L = \pi\, a\, c\, r_1^2\, T_e^4\ .$$

The radius-mass relation is of the form $r_1 \sim \mathsf{M}^m$, where the value of m is $\frac{3}{4}$ for stars of the spectral classes O to G 4; for stars of the spectral classes G 7 to M, $m = \frac{1}{2}$; for red giants, $m = 3{\cdot}4$; and for supergiants, $m = 3{\cdot}2$. The subdwarfs, subgiants and white dwarfs show no correlation of this kind. For these stars there is only a single relation, different for each group, of the form $L = f(\mathsf{M}, r_1)$. A more detailed consideration shows that the difference in the form of the relations $L = f(\mathsf{M})$ and $r_1 = f(\mathsf{M})$ for various groups of the most widely distributed stars, and also the existence of only one relation for the subdwarfs, subgiants and white dwarfs, is apparently due to differences in the spatial disposition, kinematical characteristics and age of these groups of stars. This shows that the structure of stars belonging to these groups is probably different also.

The theory of the structure of stars should yield, in accordance with the observational data, two relations of the form $L = f(\mathsf{M})$ and $r_1 = f(\mathsf{M})$ for each group of stars having such relations, and should explain how only one relation, of the form $L = f(\mathsf{M}, r_1)$, may arise in general for some stars.

It is seen from the two equations (30.23) that the two relations of the form $L = f(\mathsf{M})$ and $r_1 = f(\mathsf{M})$ can be obtained (by the successive elimination of r_1 and L from the two equations, say) only if we know the nature of the sources of stellar energy, and therefore ε, the absorption coefficient k inside the star, and the chemical composition of the star, which determines μ.

Since ε, k and μ cannot be determined directly from observation, there is some indeterminacy in the solution of the problem. Hence, in developing a correct theory of the structure of stars, investigations of the possible physical mechanisms of the generation and absorption of energy inside stars and investigations of the chemical composition of stars are extremely important. The data on the chemical composition of stellar atmospheres are given in Chapter 5. We have as yet no similar data for the internal layers of the stars, but the intensive mixing of matter which takes place in stars enables us to assume that the chemical composition of a star is homogeneous deep into its interior. Nevertheless, this does not rule out the possibility of some differences in the chemical composition of different stars, which is also indicated by the analysis of the spectra of stellar atmospheres.

If we knew the quantities ε_c and k_c, the radius and the luminosity would be unambiguously determined when the mass and the molecular weight were given (this is stated by what is called the Vogt-Russell theorem). Since observation shows us completely definite relations $L = f(\mathsf{M})$, $r_1 = f(\mathsf{M})$, we can conclude from (30.23), without any physical theory of the processes of formation and absorption of energy inside stars, that for actual stars the quantities k_c and ε_c themselves must depend in some way on at least one of the parameters r_1, L, M and μ.

The hypothesis that the density of energy sources depends only on the molecular weight (which might be the case with natural radioactivity) is in direct contradiction to the facts, and must be rejected.

The generation of energy depends directly on the mass and the radius in the case where it occurs by gravitational contraction. Let us consider this more closely. If $c_v\, T = U$ is the *internal energy* of the gas, then, by the kinetic theory of gases, the *kinetic energy* of the gas is

$$E_k = \tfrac{3}{2}\,(\gamma - 1)\, U\,,$$

where $\gamma = c_p/c_v$. Let W be the *potential energy* of a gaseous sphere (i. e. the work which must be done against gravity to disperse a gaseous sphere of radius r_1 to infinity), and $E = U + W$ the *total energy* of the gas. Next, for an assembly of particles in a steady state, we have the **virial theorem**

$$2\,E_k + W = 0 \qquad (30.24)$$

or

$$3(\gamma - 1)\,U + W = 0\,,$$

and the total energy

$$E = -(3\gamma - 4)\,U = \frac{3\gamma - 4}{3(\gamma - 1)}\,W\,. \qquad (30.25)$$

In the contraction, let the potential energy change by $\Delta W (<0)$; then the system loses the fraction $(3\gamma - 4)/3(\gamma - 1)$ of this in the form of radiation, and $\Delta U = -\Delta W/3(\gamma - 1)$ goes to heat the mass. If the emission of energy by the star is due to contraction, $L = -\Delta E/\Delta t$. If the star has contracted from infinity and has reached a radius r_1 at the time t, its potential energy is

$$W = -q\,G\,\mathrm{M}^2/r_1\,, \qquad (30.26)$$

where q is a factor depending on the structure of the star and is close to unity. Moreover,

$$L\,t = \frac{3\gamma - 4}{3(\gamma - 1)}\,q\,\frac{G\,\mathrm{M}^2}{r_1}\,, \qquad (30.27)$$

where L is the mean luminosity during the time t. From this we can estimate the time t during which a star continuously emits the energy L (what is called the **contraction time-scale of evolution**). For the Sun, it is found to be of the order of 10^7 years, and for Capella (α Aurigae) it is $\sim 10^5$ years, which is at least 10^3 times shorter than the probable duration of this star's existence. It is clear that the contraction of gaseous stars cannot be at all a long-continued source of their energy.

Hence the assumption appears most plausible that the quantities ε and k are determined by processes which depend on the physical conditions inside the star, i. e. on the density and temperature (and possibly on the chemical composition). In this case the conditions (30.22) and (30.23) give us two relations containing only L, M, r_1 and μ. The successive elimination of r_1 and M from these relations gives us two relations of the form $L = f(\mathsf{M},\mu)$, $r_1 = f(\mathsf{M},\mu)$. For stars of the main sequence, giants and supergiants, for which well-defined dependences on one parameter, $L = f(\mathsf{M})$ and $r_1 = f(\mathsf{M})$, are observed, these relations show that the molecular weight of stellar matter either is constant, or is a function of the mass alone. A more careful examination of the problem [120] shows that the assumption that the molecular weight is constant for all the stars concerned does not meet the requirements of the physical theory of the absorption and generation of energy inside stars (other than the giants and supergiants). For stars of the main sequence, agreement is achieved by assuming that the molecular weight increases towards the later spectral classes.

For subdwarf and subgiant stars, the two relations $L = f(\mathsf{M},\mu)$, $r_1 = f(\mathsf{M},\mu)$ to which we are led by the theory of stellar equilibrium can give us a single relation of the form $L = f(\mathsf{M}, r_1)$, which is in fact observed for these stars. However, a still more detailed investigation of these stars is necessary, in order to judge the influence of the chemical composition on the relation observed for them between the mass, the luminosity and the radius.

The material on white dwarfs is as yet insufficient for us to form any reliable opinion concerning the observed relations between the mass, radius and luminosity for these stars.

In order to decide the physical nature of the processes of absorption and generation of energy in the stars, it is necessary to know the conditions (i. e. the densities and temperatures) under which these processes take place. A rigorous determination of the temperature and density distribution in a star, from the surface inwards, demands the solution of the equilibrium equations, for instance the system (30.12), and this in turn requires a knowledge of the sources of energy and of the absorption coefficient inside the star, i. e. an impasse is reached. This problem can be solved by the method of successive approximations.

Chapter 31. The physical conditions inside stars

1. The temperatures and densities of stellar cores. Without any special theories of the structure of stars, we can roughly estimate the mean densities and temperatures inside stars, and this gives us an idea of the physical conditions inside stars, to a first approximation.

The mean densities, as determined by observation from the mass and radius of the star, vary within very wide limits, from 10^{-6} g/cm^3 for late-type giants to 10 or 100 g/cm^3 for main-sequence dwarfs. In the white dwarfs, the mean density reaches 10^6 g/cm^3; this shows that the matter in white dwarfs is in a peculiar (degenerate) state. The mean temperatures can be estimated from the general virial theorem (30.24), according to which the temperature is proportional to $\mu\, \mathsf{M}/r_1$. This gives us, for gaseous stars (the matter in which obeys Boyle's Law), mean temperatures between some millions and some tens of millions of degrees. These results determine the range of physical conditions within which we must study the possible physical processes of generation of energy, the nature of emission and absorption, the phase state of matter, and so on.

Our information about the temperatures and densities can be refined somewhat if we consider the structure of gaseous stars in equilibrium, with the simplest assumptions concerning the distribution of the sources of the stars' energy.

In fact, we have seen that, to solve the equations of equilibrium (30.16), we must know the dependence of the functions $\varkappa$ and ε' on the distance from the centre of the star. The dependence of the absorption coefficient on the density and temperature can be studied fairly satisfactorily. We can therefore consider, in the first approximation, two limiting cases: in the first we assume that the sources of energy are uniformly distributed (ε = constant), or almost uniformly in such a way that

$$k\, L(r)/\mathsf{M}(r) = \text{constant}\,, \tag{31.1}$$

and in the second case we assume that all the sources are concentrated at the centre of the star ($\varepsilon = 0$ everywhere except at the centre). The condition (31.1) characterises A. S. EDDINGTON's star model; a particular case of this is the star model with a constant absorption coefficient k and a constant energy source strength ε.

Let us first consider EDDINGTON's model. It is easily seen that the condition (31.1), by (30.14), can be written

$$\frac{L}{\mathsf{M}}\, k_c \frac{\varkappa\, \eta}{\psi} = \text{constant}\,,$$

which, since $\varkappa \to 1$ and $\eta/\psi \to \varepsilon_c\, \mathsf{M}/L$ at the centre ($r \to 0$), becomes $k_c\, \varepsilon_c$ = constant, and consequently the condition (31.1) can be written

$$\varkappa\, \eta/\psi = \text{constant} = \varepsilon_c/\varepsilon_m\,, \tag{31.2}$$

where $\varepsilon_m = L/\mathsf{M}$ is the mean energy source strength inside the star.

Dividing the first equation (30.16) by the second, we obtain, by (31.2),

$$\mathrm{d}\left(\sigma\tau + \frac{1-\beta_c}{\beta_c}\tau^4\right) = \frac{\varepsilon_m}{\varepsilon_c}\frac{1}{\lambda_1}\,\mathrm{d}\tau^4\,,$$

and hence, integrating, we obtain from the conditions (30.18) ($\sigma = \tau = 0$ at the boundary)

$$\sigma\tau + \frac{1-\beta_c}{\beta_c}\tau^4 = \frac{\varepsilon_m}{\lambda_1\,\varepsilon_c}\tau^4\,. \tag{31.3}$$

According to (30.18), we must at the same time have the conditions $\sigma = 1$, $\tau = 1$ fulfilled at the centre, and consequently

$$\lambda_1 = \beta_c\,\varepsilon_m/\varepsilon_c\,. \tag{31.4}$$

Substituting the expression for λ_1 from formula (30.17), we obtain

$$1 - \beta_c = k_c\,\varepsilon_c/4\pi\,c\,G\,. \tag{31.5}$$

Consequently, equation (31.3) can be written

$$\sigma = \tau^3\,,$$

and the first equation (30.16) takes the form

$$\frac{4}{\beta_c}\frac{\mathrm{d}\tau}{\mathrm{d}\xi} = -\frac{\psi}{\xi^2}\,. \tag{31.6}$$

Substituting ψ from this in the third equation (30.16), we have

$$\frac{1}{\xi^2}\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\xi^2\frac{\mathrm{d}\tau}{\mathrm{d}\xi}\right) = -\frac{\lambda_2\beta_c}{4}\tau^3\,. \tag{31.7}$$

We introduce a new variable

$$\eta = \sqrt{(\tfrac{1}{4}\lambda_2\,\beta_c)}\,\xi\,. \tag{31.8}$$

Then equation (31.7) takes the form

$$\frac{1}{\eta^2}\frac{\mathrm{d}}{\mathrm{d}\eta}\left(\eta^2\frac{\mathrm{d}\tau}{\mathrm{d}\eta}\right) = -\tau^3\,. \tag{31.9}$$

This is called Emden's equation with index 3; it is a particular case of Emden's equations

$$\frac{1}{\eta^2}\frac{\mathrm{d}}{\mathrm{d}\eta}\left(\eta^2\frac{\mathrm{d}\tau}{\mathrm{d}\eta}\right) = -\tau^n\,.$$

which determine the temperature distribution in polytropic gas spheres, i. e. gas spheres in which the gas pressure and the density at every point of the sphere are related by the equation

$$p \sim \varrho^{1+(1/n)} .$$

We require a solution of this equation which satisfies the conditions (30.18): $\tau = 1$ at the centre ($\eta = 0$) and $\tau = 0$ for some $\eta = \eta_1$. Such a solution exists for any n in the range $0 \leq n \leq 5$, and is characterised by values of η_1 and $(d\tau/d\eta)_{\eta_1}$ which are completely determinate if n is given. For $n = 3$,

$$\eta_1 = 6{\cdot}897 \ , \ \omega_0 \equiv \eta_1{}^2 (d\tau/d\eta)_{\eta_1} = -2{\cdot}018 \ . \tag{31.10}$$

These two conditions determine two more eigenvalues, α and λ_2. For at the boundary we must have $\psi = 1$, and from (31.6) and (31.8) we have

$$\psi = -\frac{4}{\beta_c} \xi^2 \frac{d\tau}{d\xi} = -\left(\frac{4}{\beta_c}\right)^{3/2} \frac{1}{\sqrt{\lambda_2}} \eta^2 \frac{d\tau}{d\eta} ,$$

and consequently the condition that $\psi = 1$ for $\eta = \eta_1$ gives us

$$\lambda_2 = (4/\beta_c)^3 \, \omega_0{}^2 , \tag{31.11}$$

or, according to the definition of λ_2 (30.17),

$$1 - \beta_c = \frac{1}{4^3 \omega_0{}^2} \frac{4 \pi G^3 a}{3 R^4} \mu^4 \beta_c{}^4 \mathsf{M}^2 . \tag{31.12}$$

This is EDDINGTON's quartic equation determining β_c from the mass M and the molecular weight μ. Furthermore, the condition $\eta = \eta_1$ (31.10) at the boundary ($\xi = 1/\alpha$) gives us, by (31.8),

$$\alpha = \sqrt{(\tfrac{1}{4} \lambda_2 \beta_c)} / \eta_1 = 4 \, \omega_0 / \eta_1 \, \beta_c , \tag{31.13}$$

if we take account of (31.11). Substituting α from (30.17), we obtain, instead of the condition (31.13),

$$T_c = \frac{\eta_1}{4 \, \omega_0} \beta_c \mu \frac{G\mathsf{M}}{R r_1} , \tag{31.14}$$

which determines the temperature at the centre of the star. The density at the centre is determined from the condition

$$\frac{R}{\mu} \varrho_c = \frac{\beta_c}{1 - \beta_c} \cdot \tfrac{1}{3} a \, T_c{}^3 . \tag{31.15}$$

Thus, instead of four conditions of the form (30.21), we have obtained, essentially, only three, (31.5), (31.12) and (31.14); the parameter λ_3

remains indeterminate, because there is no definite physical mechanism for the generation of energy in EDDINGTON's theory. In the case of a uniform distribution of energy sources, $\varepsilon_c = \varepsilon_m = L/\mathsf{M}$, and the condition (31.5) can be rewritten

$$1 - \beta_c = k_c\, L/4\,\pi\, c\, G\, \mathsf{M}\,, \tag{31.16}$$

so that the condition (31.12) will be the required relation between the mass and the luminosity of the star if k_c is known. EDDINGTON determines the value of $1 - \beta_c$ for one of the stars whose masses are well known (for instance Capella) from equation (31.12), taking $\mu = 2$. Then, from the known luminosity, he finds k_c from equation (31.16) and, taking this value of k_c to be *the same* for all stars (of the main sequence), he constructs the theoretical mass-luminosity relation [from the conditions (31.12) and (31.16)], which agrees with the observed relation for stars of the main sequence and for giants. If, however, the absorption coefficient k_c is calculated by means of the expression given by the theory, with T_c and ϱ_c obtained in EDDINGTON's model from formulae (31.14) and (31.15), the value found is 20 times less than the previous "astronomical" value. To reconcile the two values, it has been necessary to take the molecular weight $\mu \approx 1$, i. e. to assume a considerable preponderance of hydrogen (about 30 % by weight) inside the star, which agrees with the analysis of the chemical composition of stellar atmospheres. For the centre of the Sun, values of $T_c \approx 20 \times 10^6$ degrees and $\varrho_c = 70$ g/cm^3 are then obtained. The decrease of μ to 1 leads to the result that, for stars of mass less than 10 $\mathsf{M}_\odot$ (which form the majority), the part played by radiation pressure can be neglected in comparison with the gas pressure [for such stars $\beta_c \approx 1$, as is shown by calculation from formula (31.12)]. Hence, combining formulae (31.16) and (31.12), we obtain

$$L \sim \mu^4\, \mathsf{M}^3/k_c\,. \tag{31.17}$$

If we take for k_c here the expression given by a theory of absorption processes, for example Kramers' Law

$$k_c \sim \varrho_c/T_c^{\,7/2}\,, \tag{31.18}$$

then, by (31.14) to (31.16), equation (31.17) becomes one (of the necessary two) relations of the form (30.23) between the mass, the luminosity, the radius and the molecular weight [instead of EDDINGTON's simple relation between the mass and the luminosity only ($L \sim \mathsf{M}^3$)].

B. STRÖMGREN has discovered that this relation can be brought into agreement with observation by assuming that the molecular weight varies from star to star as we pass along the spectrum-luminosity diagram, owing to a change in the percentage content of hydrogen. To see this,

it is necessary to determine μ by means of (31.17) and (31.18) for a series of stars with known M, r_1 and L, i. e. to solve an equation of the form (30.22) for μ. Such a procedure shows that, if we take the helium content as zero, the hydrogen content of the stars can vary from 10 % to 50 %, lying mainly in the region of 30 %. It is seen from (31.17) that, since for actual stars a relation of the type $L \sim \mathsf{M}^3$ holds approximately, μ will change only slightly, even for considerable changes in the absorption coefficient k_c (μ^4/k_c remains approximately constant).

If now we take another limiting hypothesis and assume that all the energy sources are concentrated at the centre of the star, and that the absorption coefficient is determined by some physical law, such as (31.18), the appropriate calculations show that in this case the temperature at the centre of the Sun is of the order of 20×10^6 degrees, and the density $\varrho \approx 50$ g/cm³, i. e. almost the same values as for the model with a uniform distribution of sources. As with EDDINGTON's model, there exists in this case a very similar unique relation between M, L, r_1 and μ of the type (30.23), differing from that just considered only in the numerical values of the dimensionless parameters. Numerous investigations of star models in intermediate cases (between the case of uniform distribution of sources and that of a point source) lead to results similar to those given by these two limiting cases.

Thus we arrive at an important result: the distribution of the energy sources inside a gaseous star does not essentially affect the values of the temperature and density at the centre. The physical conditions inside the star are considerably affected only by the chemical composition of the matter and the mode of generation of energy in the star's interior.

2. **The absorption coefficient of stellar matter.** At temperatures of the order of 10^6 to 5×10^7 degrees K, the radiation of the star will be largely short-wave, similar to X-rays. The absorption of radiant energy in stars may occur by three processes: (1) the photo-ionisation of atoms of metals, which retain the K, L, . . . shells, (2) hyperbolic transitions of electrons under the action of the radiation field, in the fields of nuclei and of ionised atoms, from an orbit with a smaller energy to one with a greater energy, (3) Compton scattering of radiation by free electrons. To calculate the absorption coefficient, we must know the chemical composition of the stellar material. Owing to the high temperature, the light elements inside stars are completely ionised, and for total ionisation of atoms with atomic weight A and charge Z the mean molecular weight will be

$$\mu = A/(Z + 1) . \tag{31.19}$$

Since, for the majority of elements (apart from hydrogen and helium), $A \approx 2\,Z$, the mean molecular weight of completely ionised elements is approximately 2 (for hydrogen it is $\frac{1}{2}$, and for helium $\frac{4}{3}$).

The high temperature inside a star, which causes the high ionisation of stellar material, also has the result that the stellar gas, even at high densities, retains the properties of an *ideal gas*; since the distances between free particles (atomic residues and free electrons) are fairly large compared with the dimensions of the particles themselves, the interaction between them can be neglected.

However, the ionisation inside stars is incomplete, and the content of hydrogen and helium in the stars is very high. Let x_Z be the concentration by weight of an element with atomic number Z, so that in 1 cm^3 there are $\varrho\, x_Z$ grams of element Z. The mean number of independent particles (ions and electrons) in 1 cm^3 is $N_Z = \varrho\, x_Z/\bar{m}_Z$, where $\bar{m}_Z$ is the mean mass of a free particle of element Z, each atom of which is divided into n_Z independent particles; it is evident that

$$\bar{m}_Z = m_{\mathrm{H}}\, A_Z/n_Z\,.$$

Consequently

$$N_Z = \varrho\, x_Z\, n_Z/m_{\mathrm{H}}\, A_Z\,,$$

and the total number of independent particles is obtained by summing N_Z over all Z.

Since the gas pressure $p = N\,kT = k\,\varrho\, T/\mu\, m_{\mathrm{H}}$, the mean molecular weight

$$\mu = 1/\sum (x_Z\, n_Z/A_Z)\,. \tag{31.20}$$

Let the concentration of hydrogen by weight be X, and of helium be Y, so that the concentration of the remaining elements is $1 - X - Y$. Then, denoting $\sum n_Z/A_Z$ for the remaining elements by n_R (for hydrogen $n_Z/A_Z = 2$, and for helium $\frac{3}{4}$), it is easy to see that

$$\mu = 1/[2\,X + \tfrac{3}{4}\,Y + n_R(1 - X - Y)]\,. \tag{31.21}$$

The value of n_R is calculated from the theory of the ionisation of the K, L, M, . . . shells of the atoms with the known chemical composition.

The relative content of various elements in a star is such that, by number of atoms in unit volume, there is approximately 1000 times as much hydrogen as metals, and the relative content of the metals is distributed as follows:

K	Na	Al	Ca	Cr	Mg	Fe	Si
50	126	20	40	10	470	126	156

The estimates of the oxygen and helium contents are unreliable, but, according to contemporary data for the solar atmosphere, we should expect a high helium content (over 20 % by weight), and for stars of class B the helium content may be comparable with the hydrogen content.

The theory of the structure of stars begins, in the first approximation, from the hypothesis that the chemical composition is invariable, and takes the relative content of the metals to be the same as in Table 1 for stellar atmospheres (see Section 5.1). This makes it possible to calculate the mean molecular weight and the absorption coefficient of stellar material. Thus, a value very close to $\frac{1}{2}$ is obtained for the quantity n_R appearing in (31.21), and a somewhat smaller value if the heavy elements are more abundant inside stars than in their atmospheres. In general, the hydrogen content, followed by that of helium, has the greatest effect on the value of μ. If there is no hydrogen inside the stars, then μ is close to 2; if hydrogen is as abundant inside the stars as in their atmospheres, then μ is close to unity, since the hydrogen content is close to 35 % by weight (for example, in the Sun's atmosphere).

The effective cross-sections and the absorption coefficient for absorption processes due to bound-free and free-free transitions of electrons, and also to scattering by free electrons, are calculated from quantum mechanics for hydrogen-like atoms of metals.

The results of these calculations may be briefly formulated as follows (for details the reader should refer to S. CHANDRASEKHAR's book [28, p. 261]). Although the content of hydrogen and helium in the stars may be very considerable, the photoelectric absorption produced by the atoms of these elements does not play an important part, owing to their almost total ionisation, and the principal cause of absorption is formed by processes of *photo-ionisation of metals*; this constitutes the chief difference between the absorption processes inside stars and the continuous absorption processes in stellar atmospheres, where the photo-ionisation of hydrogen plays a leading part. However, hydrogen, helium and the nuclei and ions of light elements absorb also by means of free-free transitions. The amount of absorption due to these processes is small compared with the absorption by photo-ionisation of metals. At high temperatures, the absorption due to the scattering by free electrons becomes comparable with, and even predominates over, that due to the photo-ionisation of metals. A more detailed investigation shows that, in the density range from 0·01 to 10^4 g/cm^3, the absorption is determined mainly by the photo-ionisation of the K shells of atoms for temperatures less than 10^7 degrees K; at temperatures above about 10^8 degrees, the absorption is entirely due to scattering by free electrons. In the temperature range from 10^7 to 10^8 degrees K, the part played by scattering by electrons is the greater, the lower the density. Further, at high densities (from 10^2 g/cm^3 upwards), the absorption coefficient varies according to the law

$$k \sim 1/T^2 , \tag{31.22}$$

and is almost independent of the density ϱ. For low densities (from 0·01 to 1·0) the dependence has approximately the form

$$k \sim \varrho / T^{7/2}\,, \tag{31.23}$$

although the power of T varies considerably with the temperature, reaching its greatest value of 9/2 at temperatures of about 10^6 degrees K. At lower temperatures the absorption is due mainly to the photo-ionisation of the L shells of metal atoms, and for temperatures at which the L shell is largely complete the problem of calculating the absorption coefficient becomes extremely difficult, since in this case it is no longer possible to use the method of calculation based on the analogy with hydrogen-like atoms. Conversely, at very high temperatures, as was said above, scattering predominates, and the absorption coefficient is practically independent of the density and temperature, so that

$$k \approx \text{constant.} \tag{31.24}$$

The effect of the chemical composition on the absorption is due essentially to the hydrogen content alone, and that only at fairly high temperatures: the greater the hydrogen content, the smaller the absorption coefficient.

Thus the dependence of the absorption coefficient on the physical conditions inside a star cannot be represented by one simple analytical function [such as formula (31.23), or formulae (31.22) and (31.24), which hold only in restricted cases, when the matter is markedly degenerate and for very high temperatures respectively]. We must therefore have recourse to numerical methods.

3. **The sources of stellar energy.** Let us now consider the subatomic processes of liberation of energy which must take place at the temperatures and densities of stellar interiors considered above. As has been said, the generation of energy inside the stars cannot be explained by the process of natural radioactivity (α-decay). There is, however, another mode of energy generation, indicated by present-day nuclear physics. This consists in nuclear transformations leading to a synthesis of more complex atomic nuclei. This process is well known in nuclear physics, a classical example of it being the reaction

$$^{14}_{7}\mathrm{N} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{17}_{8}\mathrm{O} + {}^{1}_{1}\mathrm{H}:$$

the transformation of nitrogen into an isotope of oxygen by bombardment of nitrogen nuclei with α-particles. Another example of such a process is the reaction which led IRÈNE and FRÉDÉRIC JOLIOT-CURIE to the discovery of the neutron:

$$^{9}_{4}\mathrm{Be} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{12}_{6}\mathrm{C} + {}^{1}_{0}n + \gamma,$$

i.e. the transformation of a beryllium nucleus into carbon by bombardment with α-particles (here a γ-quantum and a neutron are produced). Similar transformations are also observed when lithium nuclei are bombarded by protons, e.g.

$$^7_3\text{Li} + ^1_1\text{H} \rightarrow 2\,^4_2\text{He} .$$

At the present time numerous nuclear reactions are known which are accompanied by the transformation of elements. All these reactions require the particles participating in them to have large energies (thousands or hundreds of thousands of electron-volts), which are communicated to charged particles, under the conditions existing in a terrestrial laboratory, by passing them through the electric fields produced in cyclotrons. In Nature, however, particles will acquire such energies if they are part of an extremely hot gas.

Let us consider this more closely. For nuclei to react, it is necessary that they should be able to approach one another closely. If it is necessary for the nuclei to approach to a distance a, the kinetic energy of their relative motion must be at least $Z_1 Z_2 e^2/a$, where $Z_1 e$ and $Z_2 e$ are the charges on the nuclei. The mean energy of the thermal motion of the nuclei is $3\,kT/2$, and the energy of the relative motion of the nuclei as they approach is $3\,kT$. From the condition

$$Z_1 Z_2 e^2/a \approx 3\,kT$$

(where $a \approx 10^{-11}$ cm is the upper limit to the distance at which the nuclei can react), we find

$$T \approx 5{\cdot}5 \times 10^7 Z_1 Z_2 ,$$

i.e. the temperature at which reactions can take place has its smallest value for two protons ($Z = 1$), and is then 55×10^6 degrees. At such a temperature, the majority of protons approach to a distance of approximately 10^{-11} cm, but such encounters do occur at lower temperatures also (of the order of some millions of degrees), since there is always a certain number of particles with energies above the average (the velocity distribution follows Maxwell's law). Even this rough estimate shows that, in the temperature range in which we are interested, nuclear reactions should occur, although they will be predominantly reactions between the nuclei of light elements (Z small). It is natural to ask immediately whether nuclear transformations can serve as a lasting source of the energy of the stars, and if so, whether such reactions exist which give the mean output of energy, proportional to L/M, observed for the stars. A very simple reaction between light elements is the formation of helium nuclei from hydrogen. The mass of a hydrogen atom is 1·00813

mass units, and that of a helium nucleus is 4·00386 mass units. Hence, in forming a helium nucleus from four protons, the energy evolved is

$$\Delta E = (4 \times 1{\cdot}00813 - 4{\cdot}00386)\, m_{\mathrm{H}}\, c^2,$$

and in transforming 1 gram of hydrogen into helium the energy evolved is

$$\Delta E/m_{\mathrm{H}} = 2{\cdot}57 \times 10^{19}\ \mathrm{ergs/g}\,;$$

here c is the velocity of light and m_{H} is the mass of a proton. The Sun emits approximately 2 ergs/g sec, and consequently the transformation of all the hydrogen in the Sun into helium would take about 10^{11} years. In this process the Sun's mass remains practically unchanged. We see that nuclear transformations are able to give the necessary duration of radiation from the stars, and even one which is at least two orders of magnitude greater than that found from geological data.

Let us now examine the question of which nuclear processes can give the observed output of energy; to do so, we must consider somewhat more closely the mechanism of these reactions.

For the probable number of collisions per unit time of nuclei whose energy of relative motion is in the range E to $E + \mathrm{d}E$ and of which N_1 and N_2 are contained in unit volume, kinetic theory gives the formula

$$\mathrm{d}Z = \frac{2\,N_1\,N_2}{(kT)^{3/2}}\,\sigma(E)\,\sqrt{\left(\frac{2\pi}{M}\right)}\,e^{-E/kT}\,E\,\mathrm{d}E \equiv N_1\,N_2\,\sigma(E)\,W(E)\,\mathrm{d}E, \tag{31.25}$$

where $M = m_1\,m_2/(m_1 + m_2)$, and $\sigma(E)$ is the effective collision cross-section. If the collision is accompanied by a reaction, then $\sigma(E)$ is the effective cross-section for the corresponding process, i.e. the probability of three events: the collision of the two particles (probability R^2, say), the penetration of one particle through the potential barrier surrounding the nucleus of the other particle (probability G, say), and the reaction following the incidence of the particle on the nucleus (probability Γ, e.g. the probability of β-decay or of the emission of a γ-quantum). Thus

$$\sigma(E) = R^2\,G\,\Gamma. \tag{31.26}$$

The total number of reactions in 1 cm^3 is

$$p = N_1\,N_2 \int\limits_0^\infty \sigma(E)\,W(E)\,\mathrm{d}E\,. \tag{31.27}$$

The probability of the penetration of a charged particle into the nucleus is calculated by means of the equations of wave mechanics applied to the theory of α-decay.

If we assume that the interaction potential of the two nuclei has the value

$$V(r) = Z_1 Z_2 e^2/r \quad \text{for} \quad r \geq r_0,$$
$$= V_0 \qquad \text{for} \quad r \leq r_0,$$

so that $V_0 = Z_1 Z_2 e^2/r_0$ is the "height" of the potential barrier, then calculation gives

$$G = e^{-2F}, \quad F = \frac{2\pi \sqrt{(2M)}}{h} \frac{Z_1 Z_2}{\sqrt{E}} g(x), \tag{31.28}$$

where $x = E/V_0$, and $g(x)$ is very close to $\frac{1}{2}\pi$ for energies $E \ll 10^6$ eV (the energies with which we are concerned). Substituting (31.25), (31.26) and (31.28) in (31.27) and effecting the integration, we obtain for the number of reactions in 1 g of matter

$$p = \frac{4}{3^{5/2}} \frac{\varrho x_1 x_2}{m_1 m_2} \frac{2\pi\Gamma}{h} a R^2 \exp[4\sqrt{(2R/a)}]\, \tau^2 e^{-\tau}, \tag{31.29}$$

where a is Bohr's radius

$$a = h^2/4\pi^2 M Z_1 Z_2 e^2,$$

x_1 and x_2 are the concentrations of the reacting nuclei by weight ($N = x/A\, m_{\mathrm{H}}$), and R is the total radius of the reacting particles

$$R = 1{\cdot}6 \times 10^{-13}(A_1 + A_2)^{1/3} \text{ cm}.$$

The quantity

$$\tau = 3(2\pi^4 M e^4 Z_1{}^2 Z_2{}^2/h^2 k T)^{1/3}. \tag{31.30}$$

For the number of thermonuclear reactions we can use the formula given by H. A. BETHE,

$$p = 5{\cdot}3 \times 10^{25} \varrho\, x_1 x_2 \Gamma \Phi(Z_1, Z_2)\, \tau^2 e^{-\tau}, \tag{31.31}$$

where

$$\Phi = \frac{1}{A_1 A_2 (Z_1 Z_2 A)^3} \left(\frac{8R}{a}\right)^2 e^{2\sqrt{(8R/a)}},$$

$$\tau = 42{\cdot}7\,(Z_1 Z_2)^{2/3} (A/T)^{1/3}, \; A \equiv M/m_{\mathrm{H}}. \tag{31.32}$$

The probability Γ can be determined from experimental data [28, Chapter XII]. By means of these formulae and the experimental data, BETHE has carried out a detailed study of nuclear reactions in a mixture of all the elements at temperatures of from 20 to 40 million degrees; here it was sufficient to consider only the light elements, since the greater

repulsive forces between the nuclei of the heavy elements give rise to a negligible probability of reactions involving them [see (31.28)]. It must be mentioned at once that, for the density and temperature range which we are considering, the synthesis of elements heavier than helium by nuclear reactions is impossible, since ^{5}Li and ^{5}He are unstable with respect to radioactive decay. Hence the heavy elements now existing in the stars must have been there in the early life of the stars also.

The simplest reaction is the formation of deuterium from two protons:

$$ {}^1_1\mathrm{H} + {}^1_1\mathrm{H} \to {}^2_1\mathrm{H} + e^+, \tag{31.33} $$

where e^+ is a positron. [The positron rapidly combines with an electron to give γ-quanta, so that two protons and an electron, as a result of the reaction (31.33), give a deuteron and γ-quanta.] If a star, at the time of its formation, consisted entirely of hydrogen, the reaction (31.33) would be the first stage in the synthesis of elements. The number of deuterons formed, however, is very small, as is shown by calculation, until the temperature reaches several million degrees. Furthermore, the deuterons formed enter almost instantaneously into a reaction with other protons, leading to the formation of α-particles according to the reactions

$$ \left.\begin{aligned} &{}^1_1\mathrm{H} + {}^2_1\mathrm{H} \to {}^3_2\mathrm{He} + \gamma\,, &&\text{(a)}\\ &{}^1_1\mathrm{H} + {}^3_2\mathrm{He} \to {}^4_2\mathrm{He} + e^+: &&\text{(b)} \end{aligned}\right\} \tag{31.34} $$

Calculations of the duration of the reaction (31.33) and (31.34), by means of formula (31.31), show that reaction (a) occurs 10^{18} times faster than the reaction (31.33), so that each deuteron produced will have an extremely short life, and only about one deuteron will be in statistical equilibrium with 10^{18} protons. A further combination of protons with He is impossible, since the ^{5}Li formed is unstable. The subsequent formation of ${}^4_2$He occurs much more rapidly than the reactions (31.34), according to the following chain of reactions:

$$ \left.\begin{aligned} &{}^3_2\mathrm{He} + {}^4_2\mathrm{He} \to {}^7_4\mathrm{Be} + \gamma\,, &&\text{(c)}\\ &\qquad\quad {}^7_4\mathrm{Be} \to {}^7_3\mathrm{Li} + e^+\,, &&\text{(d)}\\ &{}^1_1\mathrm{H} + {}^7_3\mathrm{Li} \to 2\,{}^4_2\mathrm{He}\,. &&\text{(e)} \end{aligned}\right\} \tag{31.34} $$

The rate of transformation of protons into α-particles by the five-membered chain (31.34) (a) to (e) depends entirely on process (c).

If p is known for the reactions, then, by calculating the change in the binding energy in the reaction, we obtain the energy from one reaction $\Delta E = \Delta m\, c^2$, and hence the energy per gram

$$ \varepsilon = p\,\Delta m\, c^2\,. $$

For the reaction considered, at a temperature of 20×10^6 degrees and a density of about 100 g/cm^3, ε is found to be about 2 ergs/g sec, if no other reactions contribute to the value of ε. Thus the formation of helium nuclei by means of the reactions (31.34) [and (31.35) below] can make a considerable contribution to the mean value of ε for the Sun (2 ergs/g sec), although the actual output of energy in the central parts of the Sun must be appreciably greater, owing to the increase of the temperature towards the centre.

It has been mentioned above that, at temperatures below a million degrees, deuterons are formed in very small amounts, and at temperatures of several million degrees, as is shown by calculations of the probability p of the reactions, the most efficient reactions are those of three light elements, lithium, beryllium and boron:

$$\left.\begin{aligned}
&{}^6_3\mathrm{Li} + {}^1_1\mathrm{H} \to {}^3_2\mathrm{He} + {}^4_2\mathrm{He};\\
&{}^7_3\mathrm{Li} + {}^1_1\mathrm{H} \to 2\,{}^4_2\mathrm{He};\\
&{}^9_4\mathrm{Be} + {}^1_1\mathrm{H} \to {}^6_3\mathrm{Li} + {}^4_2\mathrm{He};\\
&{}^{10}_5\mathrm{B} + {}^1_1\mathrm{H} \to {}^{11}_6\mathrm{C} + \gamma;\\
&{}^{11}_5\mathrm{B} + {}^1_1\mathrm{H} \to 3\,{}^4_2\mathrm{He};
\end{aligned}\right\} \qquad (31.35)$$

the last reaction begins at $T = 20 \times 10^6$ degrees. However, all these are short-period reactions (about 10^3 years), and are capable only of conserving the star's energy in the early stages of a rise in its temperature, for example by contraction; none of the reacting elements is renewed here: they are all "burnt up", in the central part of the star first of all.

After hydrogen and helium, when lithium, beryllium and boron are exhausted, carbon and nitrogen are the lightest elements; each of these has two stable isotopes. All the four nuclei can capture a proton; in two cases, the resulting nucleus is radioactively unstable. The corresponding reactions are

$$\left.\begin{aligned}
&\text{(a)}\quad {}^{12}_6\mathrm{C} + {}^1_1\mathrm{H} \to {}^{13}_7\mathrm{N} + \gamma &&\text{(2·5 million years)}\\
&\text{(b)}\quad {}^{13}_7\mathrm{N} \to {}^{13}_6\mathrm{C} + e^+ &&\text{(9·9 minutes)}\\
&\text{(c)}\quad {}^{13}_6\mathrm{C} + {}^1_1\mathrm{H} \to {}^{14}_7\mathrm{N} + \gamma &&\text{(50 thousand years)}\\
&\text{(d)}\quad {}^{14}_7\mathrm{N} + {}^1_1\mathrm{H} \to {}^{15}_8\mathrm{O} + \gamma &&\text{(4 million years)}\\
&\text{(e)}\quad {}^{15}_8\mathrm{O} \to {}^{15}_7\mathrm{N} + e^+ &&\text{(2·1 minutes)}\\
&\text{(f)}\quad {}^{15}_7\mathrm{N} + {}^1_1\mathrm{H} \to {}^{16}_8\mathrm{O} + \gamma \text{ or} &&\\
&\text{(g)}\quad {}^{15}_7\mathrm{N} + {}^1_1\mathrm{H} \to {}^{12}_6\mathrm{C} + {}^4_2\mathrm{He} &&\text{(20 years)}
\end{aligned}\right\} \qquad (31.36)$$

Consequently, if the reactions (a) to (e) lead to reaction (f), an oxygen nucleus $^{16}_{8}O$ is formed from a carbon nucleus $^{12}_{6}C$ with the emission of three γ-quanta and two positrons; if reactions (a) to (e) lead to reaction (g), the carbon nucleus is restored and the net result is that a He nucleus is formed from four protons with the emission of three γ-quanta and two positrons.

The reaction (g) is a million times more probable than (f), so that, out of a million carbon nuclei, only one is lost by reactions following the sequence (a) to (e). In all other cases the carbon nucleus is restored and takes part in a new cycle of the same reactions.

Such a cycle of reactions, in which one of the reagents is restored, is called **catalytic**: the reagent $^{12}_{6}C$ in the cycle (31.36) only "assists" in the synthesis of a helium nucleus from the protons, i. e. it is a catalyst. We may notice that all the reactions in the cycle (31.36) have been discovered in the laboratory, and their probabilities can be estimated experimentally also. The mean intervals between successive reactions, at temperatures of the order of 20×10^6 degrees K and a density of about 100 g/cm^3, are given in parentheses. Thus, the time for the whole cycle under these conditions is approximately $6{\cdot}5 \times 10^6$ years.

In order that the output of energy from this cycle of nuclear reactions (called the carbon cycle) should be of the order of 100 erg/g sec (with this energy output at the centre of the Sun, the mean energy output per gram of mass would be of the order of several ergs per gram per second), i. e. should be sufficient to conserve the Sun's energy, it is necessary that nitrogen and carbon should each form about 0·5 % (by weight) of the total composition of stellar matter. This agrees quite well with the results on the relative abundance of these elements in the Sun's atmosphere.

Similarly, we can consider the duration, under the same physical conditions, of the reactions for the capture of a proton by oxygen and by heavier elements. Calculations of the value of p, carried out by H. A. Bethe, show that for oxygen the duration of the reaction is 10^{12} years, for fluorine about 3×10^7 years, for neon approximately 2×10^{13} years, and for heavier nuclei even longer. The energy output in these reactions is in general lower than in the carbon cycle, and all the reactions involve the disappearance of elements; consequently they cannot be of importance.

Thus the source of the energy of the main-sequence stars may be the carbon cycle of nuclear reactions. In the early stages of a star's evolution, the reactions of hydrogen with the light elements Li, Be and B may be a source of energy, though a very short-lived one. Contemporary theory has not yet given a satisfactory explanation of the sources of energy in giant stars.

Chapter 32. The internal structure of stars

1. **The structure of stars of the main sequence.** We can now attempt to solve the problem of the structure of stars the source of whose energy is nuclear reactions of the type considered above, and in which the absorption is caused by the photo-ionisation of metals and scattering by free electrons. The system of equations (30.12) now contains no unknown parameters or functions: the absorption coefficient is determined, for example, by formula (31.23), or by tabulated values, for the whole range (of temperatures, densities and relative abundances of metals) in which we are interested. and the strength of the energy sources is determined, for instance, by formulae (31.31) and (31.32) for the carbon cycle of nuclear reactions. The solving of the system (30.12) in this case is a laborious process of numerical integration, which often has to be repeated until a value is found for the molecular weight μ for which the calculation, with the given mass M, gives the correct values of the luminosity L and the radius r_1 observed for the given star. It is found that such an agreement cannot be attained only by a change in the percentage content X of hydrogen; it is necessary also to take into account the content Y of helium. This circumstance is at once clear if we consider the two fundamental conditions (30.23): since k_c and ε_c in this case depend on ϱ_c and T_c, i. e. they are functions of L, M and r_1, the two conditions (30.23) are two relations between M, L and r_1 only, which in general cannot be simultaneously satisfied for a given set of three values of L, M and r_1, with only one parameter μ. The solution becomes possible only if the relations (30.23) are two relations of the form

$$F(L, \mathsf{M}, r_1, X, Y) = 0.$$

If the stars were not all formed simultaneously (a supposition which is strongly supported by observation), it is unreasonable to expect the same relative content of hydrogen and helium in stars in very different stages of evolution, if only because the conversion of hydrogen into helium is the only source of stellar energy.

The analysis of the chemical composition of stellar atmospheres shows a considerable abundance of helium there also (sometimes as much as 50%), and there is no reason to suppose that the chemical composition inside the stars differs markedly from the composition of the stellar atmospheres, since there is strong mixing of the material in stars. Moreover, calculations show that it is possible to construct a model of a main-sequence star, with the carbon cycle, only for a perfectly definite content of hydrogen and helium. The theoretical estimates of the hydrogen and helium contents thus obtained are in agreement with direct estimates of

the content of them in stellar atmospheres. For the Sun, in particular, values are obtained for the hydrogen content from 35 to 60 %, and for helium from 25 to 50 %, the content of metals being from 5 to 12 %.

Here, however, it is found that close to the centre of the star the condition of radiative equilibrium is violated, since the temperature gradient calculated from the equations of radiative equilibrium exceeds the adiabatic temperature gradient (an effect similar to the appearance of a hydrogen convective zone beneath the star's photosphere). This is due to the concentration of the energy sources close to the centre of the star, which is very marked for thermonuclear reactions, owing to the great sensitivity of the energy output to changes in temperature [see (31.29)]. The approximately steady state in such a central convective zone can be described as a state of convective equilibrium in which the system of convective currents established does not destroy the radial symmetry of the distribution of physical variables. In this case, the dependence of the pressure on the density and temperature is determined by the adiabatic equation of the stellar "gas" with some ratio of specific heats. This ratio of specific heats of the stellar gas can be obtained from the adiabatic equation

$$\mathrm{d}E + p\,\mathrm{d}v = 0,$$

where

$$p = R\,\varrho\, T/\mu + \tfrac{1}{3} a\, T^4,$$

$$E = c_v\, T + v\, a\, T^4;$$

here v is the specific volume. Substituting the values of p and E in the adiabatic equation, we have, after some elementary transformations,

$$\mathrm{d}T/T + (\Gamma_3 - 1)\mathrm{d}v/v = 0$$

or

$$\mathrm{d}p/p + \Gamma_1\, \mathrm{d}v/v = 0,,$$

where

$$\Gamma_3 = 1 + \frac{(4 - 3\beta)(\gamma - 1)}{\beta + 12(\gamma - 1)(1 - \beta)},$$

$$\Gamma_1 = \beta + \frac{(4 - 3\beta)^2(\gamma - 1)}{\beta + 12(\gamma - 1)(1 - \beta)}.$$

If $\beta = 1$ (stars of the main sequence), $\Gamma_3 = \Gamma_1 = \gamma = c_p/c_v$, the ratio of specific heats of the gas, and we have an adiabatic change of state of a single gas without radiation. If $\beta = 0$, $\Gamma_3 = \Gamma_1 = 4/3$, and we obtain the well-known expression for the adiabatic expansion of a single radiation along the adiabatic of index 4/3. Since hydrogen and helium predominate in the chemical composition of the stars, the value of γ (for $\beta = 1$) will be very close to 5/3.

The arguments given here concerning the central convective zone in the stars are, of course, very rough, since they are based on a simplified conception of adiabatic convection. In reality, the convective phenomena inside stars are considerably more complex, since they must be accompanied by heat exchange and turbulence. If these factors are taken into account, a different criterion of the onset of convection will be obtained. However, this question has hardly been investigated as yet.

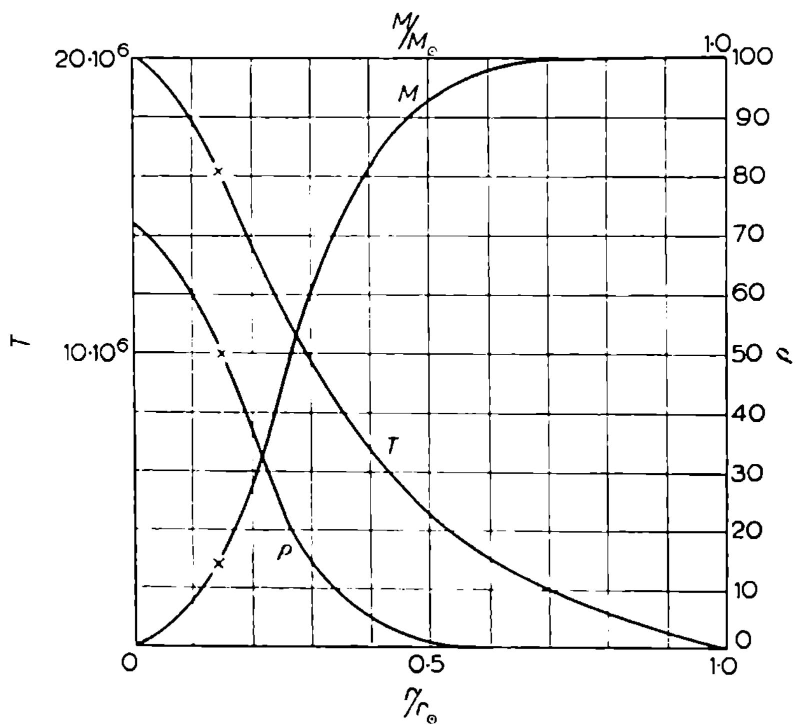

Fig. 74

In conclusion we give, as an example, the result of a numerical solution of the system (30.12) for the Sun, carried out by the method given above by A. G. Masevich and her co-workers. Table 28 shows the distance from the centre in units of the Sun's radius ($6{\cdot}96\times10^{10}$cm), the temperature (in millions of degrees), the density (in g/cm³), and finally the mass distribution

$$\psi = M(r)/M_\odot$$

and the total energy distribution $\eta = L(r)/L_\odot$. (The row marked with an asterisk gives the values at the boundary of the convective core.)

Table 28

N	$r/r_\odot$	$T\times10^{-6}$	ϱ(g/cm³)	$M(r)/M_\odot$	$L(r)/L_\odot$
1	1·000	0·006	0·000	1·000	1·00
2	0·932	0·044	0·000177	1·000	1·00
3	0·795	0·18	0·0129	1·000	1·00
4	0·676	2·20	0·110	0·993	1·00
5	0·472	5·06	1·91	0·911	1·00
6	0·358	7·78	7·84	0·743	1·00
7	0·295	9·77	15·7	0·586	1·00
8	0·262	11·00	22·0	0·482	1·00
9	0·204	13·40	36·1	0·294	1·00
10	0·169	15·03	45·5	0·189	1·00
11*	0·148*	16·11*	49·9*	0·133*	0·99*
12	0·000	20·70	71·6	0·000	0·00

It is seen from this table that all the energy is produced in a very small region at the centre (about 15% of the radius and 13% of the mass of the star). Fig. 74 shows graphically the dependence of ϱ, M and T on r.

For giant stars, as we have said, a number of difficulties arise in determining the energy sources; it is not necessary to discuss here the attempts that have been made to evolve a theory of the structure of these stars. The situation is different for the white dwarfs, the theory of whose structure is given in the next section.

2. The theory of the structure of the white dwarfs. The white dwarfs are stars with very high mean densities, up to the order of 10^6 g/cm³. Owing to the high ionisation inside the stars, the gas retains the properties of an ideal gas, obeying Boyle's Law, up to densities of about 10^2 or 10^3 g/cm³; at higher densities a degeneracy sets in, at first of the electron gas, and later (for densities greater than 10^5 or 10^6 g/cm³) of the heavy particles. In this case the gas obeys Fermi statistics. In conditions where both the electrons and the heavy particles are degenerate, the pressure of the electron gas is at least 1840 times the pressure of the heavy particles. At densities of approximately 10^6 g/cm³, relativistic effects also begin to appear in the degenerate electron gas.

Let us consider a degenerate electron gas, i. e. a gas of a density for which all the lower quantum states are occupied. All possible energy states of the electrons in a volume can be enumerated by means of quantum numbers in much the same way as are the quantum states of the electrons in an atom.

The number of quantum states with momenta between p and $p + \mathrm{d}p$ is determined by the expression

$$2\, V \cdot 4\, \pi\, p^2\, \mathrm{d}p/h^3\,,$$

where h^3 is the volume of a unit cell in momentum space, and the factor 2 takes account of the two possible spins of an electron. Next, according to Pauli's principle, only one electron can be in each quantum state. This means that, if $N(p)\,\mathrm{d}p$ is the number of electrons in a volume V whose momenta lie between p and $p + \mathrm{d}p$,

$$N\,(p)\,\mathrm{d}p \leq V \cdot 8\,\pi\, p^2\, \mathrm{d}p/h^3\,.$$

Let us consider the equation of state of a completely degenerate gas. In this case

$$N(p)\,\mathrm{d}p = V \cdot 8\,\pi\, p^2\, \mathrm{d}p/h^3\,. \tag{32.1}$$

If the volume V contains N electrons whose momenta do not exceed p_0,

$$N = V \frac{8\pi}{h^3} \int_0^{p_0} p^2 \, dp = \frac{8\pi}{3h^3} p_0{}^3 V , \tag{32.2}$$

and the electron concentration is

$$n = N/V = 8\pi p_0{}^3/3 h^3 . \tag{32.3}$$

The gas pressure p_G, by definition, is the mean value of the momentum flux through 1 cm² surface, and consequently

$$p_G V = \frac{1}{3} \int_0^{p_0} N(p)\, p\, v\, dp , \tag{32.4}$$

or, from (32.1),

$$p_G = \frac{8\pi}{3h^3} \int_0^{p_0} p^3 \frac{\partial E}{\partial p} dp , \tag{32.5}$$

where E is the kinetic energy of an electron with momentum p. From relativistic mechanics we have

$$E = m c^2 \{ \sqrt{[1 + (p^2/m^2c^2)]} - 1 \} , \tag{32.6}$$

and hence

$$\frac{\partial E}{\partial p} = \frac{p}{m} \frac{1}{\sqrt{[1 + (p^2/m^2 c^2)]}} .$$

Substituting this expression in (32.5), we obtain

$$p_G = \frac{8\pi}{3 m h^3} \int_0^{p_0} \frac{p^4 \, dp}{\sqrt{[1 + (p^2/m^2 c^2)]}} . \tag{32.7}$$

We introduce a new variable θ by the relation

$$\sinh \theta = p/m c , \ \sinh \theta_0 = p_0/m c ,$$

and equation (32.7) then reduces to

$$p_G = \frac{8\pi m^4 c^5}{3 h^3} \int_0^{\theta_0} \sinh^4 \theta \, d\theta , \tag{32.8}$$

whence. by integrating, we obtain

$$p_G = \frac{8\pi m^4 c^5}{3h^3}\left[\frac{\sinh^3\theta_0\cosh\theta_0}{4} - \frac{3\sinh 2\theta_0}{16} + \frac{3\theta_0}{8}\right]. \tag{32.9}$$

Putting

$$p_0/mc = x\,, \tag{32.10}$$

we finally obtain

$$p_G = (\pi m^4 c^5/3h^3)\, f(x) = 6{\cdot}01 \times 10^{22} f(x)\,, \tag{32.11}$$

where

$$f(x) = x(2x^2-3)\sqrt{(x^2+1)} + 3\sinh^{-1} x\,. \tag{32.12}$$

On the other hand, from (32.3) and (32.10) we have for the density of the gas the expression

$$\varrho = n\,\mu\, m_p = (8\pi m^3 c^3/3h^3)\,\mu\, m_p\, x^3\,, \tag{32.13}$$

where m_p is the mass of the proton and μ is the molecular weight. The equations (32.11) to (32.13) are the equations of state of a degenerate gas in parametric form. These equations are valid generally, and in particular for the cases of extreme non-relativistic ($p_0 \ll mc$) and extreme relativistic ($p_0 \gg mc$) degeneracy. In the former case, where x is small ($p_0 \ll mc$), we obtain, by eliminating x from (32.11) and (32.13),

$$p_G = K_1\varrho^{5/3}\,,\; K_1 = \frac{1}{20}\left(\frac{3}{\pi}\right)^{2/3}\frac{h^2}{m(\mu m_p)^{5/3}}\;(x\to 0)\,. \tag{32.14}$$

At very high densities, the value of the maximum momentum p_0 begins to exceed mc, and in the limiting case $p_0 \gg mc$ we obtain the following asymptotic expression:

$$p_G = K_2\varrho^{4/3}\,,\; K_2 = \left(\frac{3}{\pi}\right)^{1/3}\frac{hc}{8(\mu m_p)^{4/3}}\;(x\to\infty)\,. \tag{32.15}$$

Let us now consider the problem of the equilibrium of white dwarfs. To do so, we first discuss the part played by radiation pressure in a degenerate gas. Ordinary degeneracy begins when the electron pressure, calculated from formula (32.14), is considerably greater than the ordinary value of the gas pressure $p_e = n_e kT$. If we write down this inequality, we obtain a criterion for the degeneracy of the gas in the form

$$A_1 \equiv n h^3/(2\pi m kT)^{3/2} \gg 1 \tag{32.16}$$

(if factors of the order of unity are disregarded).

Let us find the ratio of the gas pressure to the radiation pressure. According to formula (32.14) we have ($p_R = \frac{1}{3} a T^4$, $a = 8\pi^5 k^4/15 c^3 h^3$)

$$\frac{p_G}{p_R} = \frac{3^{8/3} A_1^{5/3}}{\pi^{19/6}\, 2^{5/2}}\left(\frac{mc^2}{kT}\right)^{3/2}.$$

In the case of non-relativistic degeneracy, however, the mean energy of the electrons is $kT \ll m\,c^2$, and $A_1 \gg 1$; consequently, $p_G/p_R \gg 1$. Relativistic degeneracy evidently begins when the pressure (32.15) considerably exceeds the gas-kinetic pressure $n_e\,kT$. This leads to a condition of the form

$$A_2 \equiv \frac{3}{8^3\,\pi}\,\frac{n\,(h\,c)^3}{(kT)^3} \gg 1\,. \tag{32.17}$$

Forming the ratio p_G/p_R, we have by (32.15) and (32.17)

$$p_G/p_R \approx A_2^{4/3}\,,$$

whence it again follows that $p_G \gg p_R$. Consequently, radiation pressure can be completely neglected in a degenerate gas, in both the relativistic and the non-relativistic case.

Let us now consider the gravitational equilibrium of a degenerate Fermi gas. We shall take the equations of state in the most general form (32.11) to (32.13) or

$$p_G = A\,f(x)\,,\ \varrho = B\,x^3\,, \tag{32.18}$$

where

$$A = \pi\,m^4\,c^5/3\,h^3\,,\ B = 8\,\pi\,m^3\,c^3\,m_p\,\mu/3\,h^3\,, \tag{32.19}$$

and $f(x)$ is defined by formula (32.12). The equation of hydrostatic equilibrium (30.7) and the mass equation (30.9), in the case of a degenerate gas ($p_G \gg p_R$), give

$$\frac{1}{r^2}\,\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r^2}{\varrho}\,\frac{\mathrm{d}p_G}{\mathrm{d}r}\right) = -4\,\pi\,G\,\varrho\,. \tag{32.20}$$

Substituting here the expressions for p_G and ϱ from (32.18), we obtain

$$\frac{A}{B}\,\frac{1}{r^2}\,\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r^2}{x^3}\,\frac{\mathrm{d}f(x)}{\mathrm{d}r}\right) = -4\,\pi\,G\,B\,x^3\,. \tag{32.21}$$

Putting here

$$y = \sqrt{(1+x^2)} = y_0\,\Phi\,,\ r = \sqrt{(2\,A/\pi G)}\cdot\eta/B\,y_0\,, \tag{32.22}$$

and noticing that, by (32.12),

$$\frac{1}{x^3}\,\frac{\mathrm{d}f(x)}{\mathrm{d}x} = 8\,\frac{\mathrm{d}}{\mathrm{d}x}\,\sqrt{(x^2+1)}\,, \tag{32.23}$$

we obtain the fundamental equation of the problem:

$$\frac{1}{\eta^2}\,\frac{\mathrm{d}}{\mathrm{d}\eta}\left(\eta^2\,\frac{\mathrm{d}\Phi}{\mathrm{d}\eta}\right) = -\left(\Phi^2 - \frac{1}{y_0^2}\right)^{3/2}. \tag{32.24}$$

The quantities x_0 and y_0 are the values of x and y at the centre of the star, determined by the central value of the density. The solution of equation (32.24) must clearly satisfy the conditions

$$\text{for } \eta = 0\,, \Phi = 1 \text{ and } \mathrm{d}\Phi/\mathrm{d}\eta = 0\,, \tag{32.25}$$

so that the expression for the mass has the form

$$\mathsf{M}(r) = 4\,\pi \int_0^\eta \varrho\, r^2\, \mathrm{d}r = -4\,\pi \left(\frac{2\,A}{\pi\, G}\right)^{3/2} \frac{1}{B^2}\, \eta^2\, \frac{\mathrm{d}\Phi}{\mathrm{d}\eta}\,. \tag{32.26}$$

The solutions of equation (32.24) which satisfy the conditions (32.25) contain the parameter $y_0 = \surd\,(1 + {x_0}^2)$, which depends only on the density ϱ_c at the centre and on the molecular weight. Of the possible solutions, we must select the one which satisfies the conditions (30.18) (the density vanishes at the boundary, and the mass $\mathsf{M}(r) = \mathsf{M}$, the mass of the star).

Investigations have shown that, for every value of ϱ_c, there is a unique solution, i. e. the function Φ vanishes for some $\eta = \eta_1$. The second equation (32.22) (with $\eta = \eta_1$) determines the radius of the degenerate star as a function of ϱ_c and μ, i. e. it gives an equation of the form (30.22):

$$r_1 = f_1(\varrho_c, \mu)\,.$$

The second condition at the outer boundary, $\mathsf{M}(r_1) = \mathsf{M}$, gives us, by (32.26), a second relation of the form (30.22):

$$\mathsf{M} = f_2(\varrho_c, \mu).$$

Eliminating ϱ_c from these two relations, we obtain one relation of the form (30.23), i. e.

$$F(r_1, \mathsf{M}, \mu) = 0\,,$$

so that, if the mass is given (together with the molecular weight), the radius of the white dwarf is uniquely determined. Table 29 gives the values of the mass M (in units of the Sun's mass), the central density ϱ_c, the mean density ϱ_m, and the radius r_1, according to S. Chandrasekhar's calculations. (These results were obtained for $\mu = 1$; for a molecular weight other than unity, the values given in the table for M should be multiplied by μ^{-2}, those for r_1 by μ^{-1}, and those for ϱ_c and ϱ_m by μ.)

Table 29

$M/M_\odot$	Central density (g/cm³)	Mean density (g/cm³)	Radius (cm)
5·728	∞	∞	0
5·484	$9{\cdot}737 \times 10^8$	$4{\cdot}716 \times 10^7$	$4{\cdot}136 \times 10^8$
5·294	3·391	1·578	5·443
4·852	$8{\cdot}187 \times 10^7$	$5{\cdot}111 \times 10^6$	7·699
4·310	2·669	2·114	9·936
3·528	$7{\cdot}908 \times 10^6$	$7{\cdot}960 \times 10^5$	$1{\cdot}287 \times 10^9$
2·934	3·523	4·065	1·514
2·440	1·816	2·302	1·721
2·007	$9{\cdot}885 \times 10^5$	1·345	1·929
1·612	5·381	$7{\cdot}741 \times 10^4$	2·155
0·877	1·236	1·936	2·793
0	0	0	∞

We see that there cannot be gravitational equilibrium of a degenerate gas with mass

$$M > 5{\cdot}73\, M_\odot/\mu^2 \equiv M_0\,.$$

For stars of mass $M \lesssim 2{\cdot}5\, M_\odot/\mu^2$, CHANDRASEKHAR's calculations have shown that the main mass of the gas has ordinary degeneracy, while for $M \gtrsim 2{\cdot}5\, M_\odot/\mu^2$ relativistic degeneracy plays the chief part. The radius of the equilibrium configuration continually decreases as the mass increases, and becomes vanishingly small as the mass approaches the critical value M_0. Although the observational data on the masses and radii of the white dwarfs are not plentiful, they show that the values given in the table are of the same order as those observed.

We see, however, that, in solving the problem of the equilibrium of the white dwarfs, the question of their luminosities has remained in abeyance. It cannot be solved unless we know the sources of energy and the molecular weight inside the star. For the white dwarfs we as yet have no reliable data on which to base an account of their luminosity. It is possible that the gravitational contraction of such dense stars may be the source of their energy, but the problem of the sources of the energy of white dwarfs still requires detailed study.

PART VIII.
THE SCATTERING OF LIGHT IN PLANETARY ATMOSPHERES

Chapter 33. The theory of radiative transfer in planetary atmospheres

1. **Radiative transfer in planetary atmospheres. The indicatrix of scattering.** It is well known that the planets Venus, Jupiter and Saturn, and also apparently Uranus and Neptune, are surrounded by atmospheres so dense that we do not see the surfaces of the planets. The radiation reaching us from these planets is the Sun's radiation scattered by their atmospheres. Hence the laws obeyed by the light reflected from the planets (the distribution of brightness over the disc, the change in the total brightness with phase, the numerical value of the albedo) can be obtained only from a study of the processes of scattering in planetary atmospheres. Since the atmospheres of these planets are opaque, i. e. they have a very large optical thickness ($\tau \gg 1$), a light quantum entering the atmosphere of a planet has a fairly large probability of undergoing, before leaving the atmosphere, not one but many scattering processes, i. e. we are essentially concerned with multiple scattering of light in the atmospheres of the planets.

The study of the laws of optics in the Earth's atmosphere has also led to a consideration of scattering processes. LEONARDO DA VINCI explained the blue colour of the sky as being due to the scattering of sunlight by the air and the more intensive scattering of blue light than of red light. It is true that the optical thickness of the Earth's atmosphere is, in the absence of cloud, considerably less than unity. However, numerous investigations have shown that, even in explaining the distribution of brightness over the sky, scatterings of orders higher than the first play a certain part. (We are not speaking of the case where a layer of cloud is present, when the sunlight undergoes a large number of scatterings in this layer before reaching the surface of the Earth.) Thus, the problem of the scattering of light in the Earth's atmosphere is also, in general, one of multiple scattering in a medium. The problem of the multiple scattering of light in planetary atmospheres is solved by the study of radiative transfer. Here, since the geometrical thickness

of the atmosphere is in every case small in comparison with the radius of curvature of the atmospheric layers (this is always valid for the thickness of that part of a planet's atmosphere from which any appreciable fraction of the quanta scattered towards us actually reaches us directly), we can assume in our calculations that the atmosphere consists of plane-parallel layers.

If processes of *pure scattering* alone took place in a planetary atmosphere, then, for a sufficiently great optical thickness, it would totally reflect all the energy incident on it. In other words, the albedo of the planet would be unity in this case. However, observation shows that the albedo of many planets having opaque atmospheres is considerably less than unity. It follows from this that the energy absorbed in the visible and photographic wavelengths by some element of volume of the planetary atmosphere is not completely scattered by it, but is partly transformed into other forms of energy, such as thermal energy. In accordance with the usually accepted notation in the theory of planetary atmospheres, we shall here denote by λ *the fraction of the energy in a given frequency that is scattered immediately, after being absorbed in a single process*, in the same frequency. Then $1 - \lambda$ is the fraction of the energy which, after being absorbed in a single process, is converted into other forms of energy, and is not re-emitted immediately in the same frequency, i. e. undergoes *true absorption.* (In the theory of stellar photospheres, this quantity is denoted by ε_ν.)

It is evident that each element of volume in the atmosphere not only causes scattering of the light passing through it, but is also a source of thermal radiation, since the temperature of the atmosphere must be fairly high, because of the processes (mentioned above) of the partial conversion of the absorbed energy into heat (true absorption). However, it can be shown that the thermal radiation of planetary atmospheres, at the temperatures existing in them, must be mainly in the far infra-red region of the spectrum, where the wavelength exceeds 40,000 Å. The thermal radiation in the visible and photographic parts of the spectrum must be negligibly small at the temperatures existing in planetary atmospheres. Hence, although we find true absorption here ($\lambda < 1$), *the true emission in visible and photographic light can be neglected.* Thus, in each of these frequencies, each element of volume scatters only some definite fraction of the absorbed energy, and emits nothing in addition. This means that, since we are concerned with the visible and photographic parts of the spectrum, the radiation field in each frequency can be considered independently of the radiation field in other frequencies. If the energy scattered by each element of volume in a single process were distributed isotropically, independently of the direction in which it was propagated before this process (in this case

we say that the element of volume has a *spherical indicatrix of scattering*), the condition of radiative equilibrium could be written in the following form:

$$4\,\pi\, j_\nu = \lambda\, \alpha_\nu \int I_\nu\, \mathrm{d}\omega\,, \tag{33.1}$$

where $4\,\pi\, j_\nu$ is the energy emitted in all directions by unit volume in unit frequency interval, α_ν is the coefficient of attenuation of light in the given frequency or coefficient of extinction, I_ν is the intensity of radiation, and $\mathrm{d}\omega$ is an element of solid angle.

The quantity $s_\nu = \lambda\, \alpha_\nu$ is often called the scattering coefficient, and the quantity $k_\nu = (1 - \lambda)\, \alpha_\nu$ is the true absorption coefficient. In general, the value of λ depends on the frequency ν.

A quantum going in any given direction has in general, after scattering, a different probability of being emitted in different directions. We denote by θ and ϕ the angles defining the direction of motion of the quantum before scattering, in some system of spherical co-ordinates. After scattering, the quantum is emitted in a direction defined by the angles θ' and ϕ', say. We take an element of solid angle $\mathrm{d}\omega'$ about the direction (θ', ϕ'). What is the probability that, after scattering, the quantum has a direction lying within some $\mathrm{d}\omega'$? It is evident that this probability is proportional to $\mathrm{d}\omega'$, and must depend, in general, on the angle between the original direction (θ, ϕ) of the quantum and its final direction (θ', ϕ'). We denote this angle by γ. Then

$$\cos\gamma = \cos\theta\,\cos\theta' + \sin\theta\,\sin\theta'\,\cos(\phi - \phi')\,. \tag{33.2}$$

For the required probability we obtain the expression

$$\mathrm{d}P = x(\cos\gamma)\,\mathrm{d}\omega'/4\,\pi\,. \tag{33.3}$$

The function $x(\cos\gamma)$ is called the indicatrix of scattering or the *phase function* characterising the scattering process. Since the sum of the probability over all directions must be equal to unity, we have from (33.3)

$$\int x(\cos\gamma)\,\mathrm{d}\omega'/4\,\pi = 1\,. \tag{33.4}$$

In the case mentioned above, where the probability of scattering in any direction is the same (a spherical indicatrix of scattering), we must have $x(\cos\gamma) =$ constant, and we then have immediately, from (33.4),

$$x(\cos\gamma) = 1\,. \tag{33.5}$$

In general, $x(\cos\gamma)$ is some positive function of $\cos\gamma$, which can be expanded in a series of Legendre polynomials:

$$x(\cos\gamma) = x_0 + x_1\,P_1(\cos\gamma) + x_2\,P_2(\cos\gamma) + \ldots\,. \tag{33.6}$$

Here it follows directly from (33.4) that

$$x_0 = 1 . \tag{33.7}$$

For a spherical indicatrix of scattering, the expansion (33.6) is evidently limited to one term.

If the indicatrix of scattering is given, the condition of radiative equilibrium can easily be written down under the conditions described above. The energy emitted in unit frequency interval by unit volume in the solid angle $\mathrm{d}\omega'$ in unit time must be

$$j_\nu \, \mathrm{d}\omega' = \lambda \, \alpha_\nu \int x(\cos\gamma) \, I_\nu(\theta, \phi) \frac{\mathrm{d}\omega}{4\pi} \mathrm{d}\omega' , \tag{33.8}$$

since, out of each element of solid angle $\mathrm{d}\omega$, an amount of energy $\alpha_\nu I_\nu(\theta, \phi)\,\mathrm{d}\omega$ is absorbed by unit volume; of this, $\lambda \, \alpha_\nu I_\nu(\theta, \phi)\,\mathrm{d}\omega$ is scattered, and to find the fraction of this scattered inside $\mathrm{d}\omega'$ we must multiply this quantity by $x(\cos\gamma)\,\mathrm{d}\omega'/4\pi$, in accordance with (33.3).

Dividing (33.8) by $\mathrm{d}\omega'$, we obtain the condition of monochromatic radiative equilibrium in the form

$$j_\nu = \lambda \, \alpha_\nu \int x(\cos\gamma) \, I_\nu(\theta, \phi) \, \mathrm{d}\omega / 4\pi . \tag{33.9}$$

This equation must be considered together with the equation of transfer, which, in the case of plane-parallel layers with the angles θ and θ' measured from the outward normal, has the form

$$\cos\theta' \, \mathrm{d}I_\nu/\mathrm{d}z = - \alpha_\nu I_\nu + j_\nu , \tag{33.10}$$

where j_ν depends not only on the height z in the atmosphere, but also on the direction (θ', ϕ').

Substituting for j_ν the expression (33.9) and introducing the optical depth τ_ν in the frequency concerned, we obtain the integro-differential equation

$$\cos\theta' \, \mathrm{d}I_\nu/\mathrm{d}\tau_\nu = I_\nu - \lambda \int x(\cos\gamma) \, I_\nu(\theta, \phi) \, \mathrm{d}\omega/4\pi , \tag{33.11}$$

which contains only one unknown function I_ν, depending on τ_ν, θ and ϕ.

We shall introduce the following simplification of the notation. Primes denote angular variables over which we integrate; those over which we do not integrate will be written without primes. The suffix ν will generally be omitted. Since the radiation field is calculated independently for each frequency, this cannot lead to any misunderstanding.

Equation (33.11) is then written in the form

$$\cos\theta\,\mathrm{d}I/\mathrm{d}\tau = I - \lambda \int x(\cos\gamma)\, I(\theta',\phi')\,\mathrm{d}\omega'/4\pi\,, \qquad (33.12)$$

where $\cos\gamma$ is defined by (33.2).

For the problem here considered, we have, besides equation (33.12), the following boundary condition. The intensity of the incident radiation is given at the outer boundary of the atmosphere ($\tau = 0$), and it is zero for all directions except those in which the direct rays of the Sun are incident.

The flux of direct solar rays, measured by the quantity of energy incident on unit area perpendicular to the rays in unit time and unit frequency interval, is denoted by πS. Since, in this case, we are concerned with radiation concentrated in one direction, it is convenient, in order to avoid the use of discontinuous functions, to exclude the direct radiation of the Sun from the intensity of radiation I, introduced above, and to understand by I only the intensity in the field of diffuse radiation scattered by the medium itself. Then the total intensity of radiation is equal to the sum of the scattered radiation I and some function which differs from zero for all directions, but is infinitely large in some direction (θ_0, ϕ_0), namely that of the direct rays of the Sun.

As a result of this, both these components must appear in the integral on the right-hand side of (33.12), and since we have retained the notation I only for the intensity of the diffuse radiation (the first component), we must write instead of this integral the sum

$$\lambda \int x(\cos\gamma)\; I(\theta',\phi')\,\mathrm{d}\omega'/4\pi + \lambda(\pi S/4\pi)\; x(\cos\gamma_0)\; e^{\tau\sec\theta_0}, \quad (33.13)$$

where γ_0 is the angle between the directions (θ, ϕ) and (θ_0, ϕ_0), since the total intensity of sunlight outside the atmosphere is πS (i.e. the integral of the intensity of the direct solar rays over the solid angle subtended by the Sun's disc is πS), but at an optical depth τ this intensity, owing to the attenuation (extinction) of the direct rays, is reduced to $\pi S\, e^{\tau\sec\theta_0}$. Since θ_0 is the angle to the outward normal to the atmospheric layers at which the Sun's rays are incident, $\sec\theta_0$ is negative.

Under these conditions regarding the meaning of the function $I(\theta, \phi)$, we must write, instead of equation (33.12),

$$\cos\theta\,\mathrm{d}I/\mathrm{d}\tau = I - \lambda \int x(\cos\gamma)\; I(\theta',\phi')\,\mathrm{d}\omega'/4\pi - \tfrac{1}{4}\lambda S\; x(\cos\gamma_0)\; e^{\tau\sec\theta_0}, \qquad (33.14)$$

where we have for I the simple boundary condition $I = 0$ for $\tau = 0$ and $\theta > \frac{1}{2}\pi$, i.e. the intensity of incident diffuse radiation at the boundary of the atmosphere is zero.

By solving equation (33.14), we find the value of $I(\theta',\phi')$ for all τ, i.e. for all depths in the atmosphere and, in particular, for $\tau = 0$. By finding the value of $I(\theta',\phi')$ for $\tau = 0$, we determine the distribution of brightness over the planet's disc. The results obtained can then be directly compared with observation.

2. **The principle of invariance and its application.** The integro-differential equation (33.14) is usually replaced by the integral equation obtained from it by using the boundary condition; the unknown function is

$$B(\tau,\theta,\phi) = \lambda \int x(\cos\gamma)\, I(\theta',\phi')\, d\omega'/4\pi + \tfrac{1}{4}\lambda S\, x(\cos\gamma_0)\, e^{\tau \sec\theta_0}. \quad (33.15)$$

If, by solving this integral equation, the function $B(\tau,\theta,\phi)$ is found, then equation (33.14) is converted into an ordinary inhomogeneous linear differential equation soluble by quadratures, and, in particular, the value of I for outward directions ($\theta < \frac{1}{2}\pi$) at $\tau = 0$ is found. The value of B determines the emission coefficient j, since, according to (33.9) and (33.15),

$$j = \alpha B(\tau,\theta,\phi)\,. \quad (33.16)$$

Thus it is found that, to determine I even at the boundary and to compare this result with observation, it is necessary to determine the value of the unknown function $B(\tau,\theta,\phi)$ or j at all depths and for all θ and ϕ. Hence, if we are interested only in the distribution of brightness over the planet's disc, this method is extremely clumsy.

In short, whilst we are interested only in some function of two variables, i.e. the values of $I(\theta,\phi)$ for $\tau = 0$, the method of reducing the problem to an integral equation compels us to find a function $B(\tau,\theta,\phi)$ of three variables. The question arises whether one can construct an equation from which the values of $I(\theta,\phi)$ at the boundary $\tau = 0$ would be obtained directly.

It is found that such an equation can be constructed. Only $I(0,\theta,\phi)$, i.e. the angular distribution of the emergent radiation, will appear in it as an unknown function. Moreover, the equation obtained reveals at once (without a detailed solution) the structure of the function $I(0,\theta,\phi)$.

To derive the required equation, we shall begin from the *principle of invariance*, which has been widely used in recent years in problems of the theory of scattering of light. The problem under discussion amounts to the following. A flux of parallel rays in a direction defined by the angles (θ_0,ϕ_0) is incident on the outer boundary of a medium of infinite optical thickness. As a result of scattering processes occurring in the medium, some fraction of the rays are, as it were, reflected by the medium, this *reflection* taking place *in all directions*. Such reflection is called **diffuse reflection.** We require the intensity of the diffusely reflected

radiation. It is evident that the intensity of the diffusely reflected light is a function both of the angles of incidence θ_0, ϕ_0, and of the angles of reflection θ, ϕ. We can therefore write

$$I = I(\theta, \phi; \theta_0, \phi_0) .$$

It is also evident that the reflected intensity is proportional to that of the incident radiation, i.e. to the quantity S. Hence

$$I(\theta, \phi; \theta_0, \phi_0) = S\, r(\theta, \phi; \theta_0, \phi_0) . \tag{33.17}$$

The quantity $r(\theta, \phi; \theta_0, \phi_0)$ may be called the **diffuse reflection coefficient.** Unlike what has been done above, we shall from now on measure the angle θ_0 from the inward normal, since this is more convenient and cannot lead to any misunderstandings. The angle θ will be measured from the outward normal as before.

The principle of invariance, as applied to the problem in question, consists in the following: the diffuse reflecting power of a medium of infinite optical thickness, consisting of plane-parallel layers, cannot be changed if we add to the outer boundary a plane layer of finite optical thickness whose elements have the same optical properties as those of the original medium. The latter condition means that the coefficient λ and the indicatrix of scattering in the added layer are the same as in the original medium. Here, for simplicity, we shall first consider the case of a spherical indicatrix of scattering, where $x(\cos\gamma) = 1$.

We denote the boundary of the medium, before the addition of the layer, by A, and the new boundary, after the addition of the layer, by A'. The optical thickness of the added layer we denote by $\Delta\tau$. Since the choice of $\Delta\tau$ is at our disposal, we shall take it so small that its square can be neglected in comparison with $\Delta\tau$ itself. We shall find the changes in the intensity of the radiation diffusely reflected from the medium which occur when the layer $\Delta\tau$ is added, and equate the algebraic sum of these changes to zero.

Let us enumerate these changes:

(1) The radiation from the Sun, in passing through the additional layer, is attenuated by a factor $1 - (\Delta\tau/\cos\theta_0)$. After being diffusely reflected from the boundary A it is again attenuated, by a factor $1 - (\Delta\tau/\cos\theta)$. Hence, instead of the intensity $I(\theta, \theta_0)$, an intensity

$$I(\theta, \theta_0)\left(1 - \frac{\Delta\tau}{\cos\theta_0}\right)\left(1 - \frac{\Delta\tau}{\cos\theta}\right)$$

emerges.

We notice that, in consequence of the spherical indicatrix of scattering, the intensity of the reflected light does not depend on the azimuthal angles of the incident and emergent rays, and consequently, instead of $I(\theta, \phi; \theta_0, \phi_0)$, we write simply $I(\theta, \theta_0)$.

Thus, as a result of the attenuation of the incident and diffusely reflected rays in the additional layer, the intensity of the diffusely reflected radiation is diminished by

$$I(\theta, \theta_0)\, \Delta\tau \left(\frac{1}{\cos\theta} + \frac{1}{\cos\theta_0}\right).$$

(2) The additional layer $\Delta\tau$ scatters directly into the direction θ a part of the direct radiation of the Sun incident on it. On account of this intensity, the diffusely reflected light is increased by

$$\frac{\lambda}{4\pi}\, \pi S \frac{\Delta\tau}{\cos\theta}.$$

(3) The additional layer $\Delta\tau$, in scattering the direct rays of the Sun incident on it, directs part of the energy of these rays on to the boundary A at different angles θ' to the normal. They undergo diffuse reflection from the boundary A; as a result, additional radiation is obtained of intensity

$$\int_0^\pi d\theta' \int_0^{2\pi} d\phi'\, I(\theta, \theta') \frac{\sin\theta'}{\cos\theta'} \frac{\lambda}{4\pi} \Delta\tau,$$

which, after integrating over ϕ', gives

$$\tfrac{1}{2}\lambda \Delta\tau \int_0^\pi I(\theta, \theta') \frac{\sin\theta'}{\cos\theta'}\, d\theta'.$$

(4) The additional layer, in absorbing some fraction of the light diffusely reflected from the boundary A, scatters part of this light in the given direction θ. The intensity of this radiation is

$$\frac{\lambda}{4\pi} \frac{\Delta\tau}{\cos\theta} \int_0^\pi d\theta' \int_0^{2\pi} d\phi'\, I(\theta', \theta_0) \sin\theta' = \tfrac{1}{2}\lambda \frac{\Delta\tau}{\cos\theta} \int_0^\pi I(\theta', \theta_0) \sin\theta'\, d\theta'.$$

(5) The direct solar rays are diffusely reflected from the boundary A, are scattered by the additional layer $\Delta\tau$ back in the direction of A, and are again diffusely reflected from the surface A. This gives an additional radiation

$$\Delta\tau \int_0^\pi \int_0^{2\pi} \frac{I(\theta, \theta'')}{\pi S} \frac{\sin\theta''\, d\theta''\, d\phi''}{\cos\theta''} \frac{\lambda}{4\pi} \int_0^\pi d\theta' \int_0^{2\pi} d\phi'\, I(\theta', \theta_0) \sin\theta' =$$

$$= \lambda \Delta\tau \int_0^\pi \frac{I(\theta, \theta'')}{S} \frac{\sin\theta''}{\cos\theta''}\, d\theta'' \int_0^\pi I(\theta', \theta_0) \sin\theta'\, d\theta'.$$

All remaining additions to the radiation are of the second or higher order in $\Delta\tau$.

The last four contributions, which are positive, must balance the first contribution, which is negative, in order that the intensity of the diffusely reflected light should remain unchanged despite the addition of a layer. Hence, neglecting quantities of the second order in $\Delta\tau$, we find

$$I(\theta,\theta_0)\,\Delta\tau\left(\frac{1}{\cos\theta}+\frac{1}{\cos\theta_0}\right)=\tfrac{1}{4}\,\lambda\,S\,\frac{\Delta\tau}{\cos\theta}+$$
$$+\tfrac{1}{2}\,\lambda\,\Delta\tau\int_0^\pi I(\theta,\theta')\frac{\sin\theta'\,\mathrm{d}\theta'}{\cos\theta'}+\tfrac{1}{2}\,\lambda\,\frac{\Delta\tau}{\cos\theta}\int_0^\pi I(\theta',\theta_0)\sin\theta'\,\mathrm{d}\theta'+$$
$$+\frac{\lambda\,\Delta\tau}{S}\int_0^\pi I(\theta,\theta'')\frac{\sin\theta''}{\cos\theta''}\,\mathrm{d}\theta''\int_0^\pi I(\theta',\theta_0)\sin\theta'\,\mathrm{d}\theta'\,.$$

We divide both sides of this equation by $\Delta\tau$, and introduce, instead of the angles θ and θ_0, their cosines:

$$\cos\theta=\eta\,,\quad \cos\theta_0=\xi\,,\quad \cos\theta'=\eta'\,. \tag{33.18}$$

Then we find that the condition of invariance gives

$$I(\eta,\xi)\left(\frac{1}{\eta}+\frac{1}{\xi}\right)=\tfrac{1}{4}\,\lambda\,S/\eta+\tfrac{1}{2}\,\lambda\int_0^1 I(\eta,\eta')\,\mathrm{d}\eta'/\eta'+$$
$$+\frac{\lambda}{2\eta}\int_0^1 I(\eta',\xi)\,\mathrm{d}\eta'+\frac{\lambda}{S}\int_0^1 I(\eta,\eta'')\frac{\mathrm{d}\eta''}{\eta''}\int_0^1 I(\eta',\xi)\,\mathrm{d}\eta'\,. \tag{33.19}$$

Finally, we introduce in this equation, instead of $I(\eta,\xi)$, the function $r(\eta,\xi)$ according to equation (33.17), i.e. we put

$$I(\eta,\xi)=S\,r(\eta,\xi)\,. \tag{33.20}$$

We then have the functional equation for $r(\eta,\xi)$

$$r(\eta,\xi)\left(\frac{1}{\eta}+\frac{1}{\xi}\right)=\tfrac{1}{4}\,\lambda/\eta+\tfrac{1}{2}\,\lambda\int_0^1 r(\eta,\eta')\,\mathrm{d}\eta'/\eta'+$$
$$+\frac{\lambda}{2\eta}\int_0^1 r(\eta',\xi)\,\mathrm{d}\eta'+\lambda\int_0^1 r(\eta,\eta'')\frac{\mathrm{d}\eta''}{\eta''}\int_0^1 r(\eta',\xi)\,\mathrm{d}\eta'\,, \tag{33.21}$$

which can also be written in the form

$$r(\eta,\xi)\left(\frac{1}{\eta}+\frac{1}{\xi}\right)=\frac{\lambda}{4\eta}\left[1+2\,\eta\int_0^1 r(\eta,\eta')\,\mathrm{d}\eta'/\eta'\right]\left[1+2\int_0^1 r(\eta',\xi)\,\mathrm{d}\eta'\right]. \tag{33.22}$$

If we write also

$$r(\eta,\xi)=\xi\,\varrho(\eta,\xi)\,, \tag{33.23}$$

then we have for the function $\varrho(\eta,\xi)$, usually called the **brightness coefficient** of the reflecting layer,

$$\varrho(\eta,\xi)\,(\eta+\xi)=\tfrac{1}{4}\,\lambda\left[1+2\,\eta\int_0^1 \varrho(\eta,\eta')\,\mathrm{d}\eta'\right]\left[1+2\,\xi\int_0^1 \varrho(\eta',\xi)\,\mathrm{d}\eta'\right]. \tag{33.24}$$

It is easy to see that, if equation (33.24) is satisfied by some function $\varrho(\eta,\xi)$, it must also be satisfied by the function $\varrho(\xi,\eta)$. However, since equation (33.24) can have, from its physical significance, only one solution, the function $\varrho(\eta,\xi)$ must be symmetrical:

$$\varrho(\eta,\xi)=\varrho(\xi,\eta)\,. \tag{33.25}$$

In this case, however, it is evident that the right-hand side of (33.24) is the product of some function of η and the same function of ξ. We denote these functions by $\phi(\eta)$ and $\phi(\xi)$ respectively. Thus

$$\phi(\eta)=1+2\,\eta\int_0^1 \varrho(\eta,\eta')\,\mathrm{d}\eta'\,, \tag{33.26}$$

and

$$\varrho(\eta,\xi)=\tfrac{1}{4}\,\lambda\,\frac{\phi(\eta)\,\phi(\xi)}{\eta+\xi}\,. \tag{33.27}$$

Formula (33.27), though it does not give an explicit expression for the function $\varrho(\eta,\xi)$, reveals the structure of this function.

If we can determine the auxiliary function $\phi(\eta)$, we shall thereby find the brightness coefficient $\varrho(\eta,\xi)$ also. To obtain an equation for $\phi(\eta)$, we substitute (33.27) in (33.26), which gives

$$\phi(\eta)=1+\tfrac{1}{2}\,\lambda\,\eta\int_0^1 \frac{\phi(\eta)\,\phi(\xi)}{\eta+\xi}\,\mathrm{d}\xi\,. \tag{33.28}$$

This is a functional equation for the function $\phi(\eta)$. Thus the solution of the functional equation (33.24) has been reduced to that of the simpler functional equation (33.28), which is easily effected numerically. In

doing so, it is most convenient to use the method of successive approximations, which consists in substituting an approximate value of the function $\phi(\eta)$ on the right-hand side; the left-hand side then gives the next approximation. As an initial zero-order approximation it is best to take the mean value of $\phi(\eta)$ in the interval (0,1), i.e.

$$\phi_0 = \int_0^1 \phi(\eta)\,\mathrm{d}\eta\,.$$

The exact value of this integral is easily obtained. To do so, we integrate (33.28) over η:

$$\int_0^1 \phi(\eta)\,\mathrm{d}\eta = 1 + \tfrac{1}{2}\lambda \int_0^1\int_0^1 \frac{\phi(\eta)\,\phi(\xi)\,\eta}{\eta+\xi}\,\mathrm{d}\eta\,\mathrm{d}\xi =$$

$$1 + \tfrac{1}{4}\lambda \int_0^1\int_0^1 \frac{\phi(\eta)\,\phi(\xi)\,\eta}{\eta+\xi}\,\mathrm{d}\eta\,\mathrm{d}\xi + \tfrac{1}{4}\lambda \int_0^1\int_0^1 \frac{\phi(\eta)\,\phi(\xi)\,\xi}{\eta+\xi}\,\mathrm{d}\eta\,\mathrm{d}\xi\,,$$

since the last two integrals are equal, differing only in the naming of the variables of integration. Adding these integrals, we obtain

$$\int_0^1 \phi(\eta)\,\mathrm{d}\eta = 1 + \tfrac{1}{4}\lambda \int_0^1 \phi(\eta)\,\mathrm{d}\eta \int_0^1 \phi(\xi)\,\mathrm{d}\xi\,,$$

or

$$\phi_0 = 1 + \tfrac{1}{4}\lambda\,\phi_0^2\,,$$

whence

$$\phi_0 \equiv \int_0^1 \phi(\eta)\,\mathrm{d}\eta = [2 - 2\sqrt{(1-\lambda)}]/\lambda\,.$$

Substituting this value instead of $\phi(\eta)$ on the right-hand side of (33.28), we obtain the next approximation for $\phi(\eta)$, and so on.

For example, in the first approximation we have

$$\phi(\eta) = 1 + \tfrac{1}{2}\lambda\,\phi_0^2(\lambda)\,\eta\,\log_e[(1+\eta)/\eta]\,.$$

Table 30 gives the numerical values of the auxiliary function $\phi(\eta)$ for $\lambda = 0{\cdot}4, 0{\cdot}5, \ldots, 1{\cdot}0$. For small values of λ, it is sufficient to use the first approximation, and the numerical values of $\phi(\eta)$ in this range are therefore not given.

Table 30

η \ λ	0·4	0·5	0·6	0·7	0·8	0·9	1·0
0·0	1·00	1·00	1·00	1·00	1·00	1·00	1·00
0·1	1·06	1·07	1·09	1·11	1·14	1·17	1·25
0·2	1·09	1·11	1·15	1·18	1·23	1·29	1·45
0·3	1·11	1·14	1·19	1·24	1·30	1·39	1·64
0·4	1·13	1·17	1·22	1·28	1·36	1·48	1·83
0·5	1·14	1·19	1·25	1·32	1·41	1·56	2·01
0·6	1·15	1·20	1·27	1·35	1·46	1·63	2·19
0·7	1·16	1·22	1·29	1·38	1·50	1·69	2·38
0·8	1·17	1·23	1·31	1·40	1·54	1·75	2·55
0·9	1·18	1·24	1·32	1·42	1·57	1·80	2·73
1·0	1·18	1·25	1·34	1·44	1·60	1·85	2·91

Table 30 is to be used in the following manner. From (33.20), (33.23) and (33.27), we have, in the case of a spherical indicatrix of scattering,

$$I(\eta, \xi) = \tfrac{1}{4}\, \lambda\, S\, \xi\, \frac{\phi(\eta)\, \phi(\xi)}{\eta + \xi}\,. \tag{33.29}$$

For each point of the planet's disc, the cosine ξ of the angle of incidence of the Sun's rays is calculated, together with the cosine η of the zenith angle at which the Earth is seen from this point on the planet. The value of S can be determined, since the distance of the planet from the Sun is always known. Taking some λ, we can then calculate the intensity I and compare it with observation. If, for some λ, the values calculated from formula (33.29) and Table 30 are in fairly good agreement with the brightnesses of the various points on the planet's disc, we may conclude that the hypothesis of a spherical indicatrix of scattering is valid, at least within the limits of observational accuracy.

We notice that, for the remote outer planets, beginning with Jupiter, it can be assumed that the angles θ and θ_0 are equal, i. e. $\eta = \xi$, not only at opposition, but also at a considerable distance from opposition. In this case, formula (33.29) reduces to

$$I(\eta) = \tfrac{1}{8}\, \lambda\, S[\phi(\eta)]^2\,, \tag{33.30}$$

and we see that the greatest contrast theoretically possible between the centre of the disc ($\eta = 1$) and the limb ($\eta = 0$) is for $\lambda = 1$, when

$$I(1)/I(0) = 8{\cdot}41\,.$$

In reality, the contrasts observed for the outer planets are not so great; this is due to the fact that $\lambda < 1$.

3. Generalisation to the case of a non-spherical indicatrix of scattering. In the case of a non-spherical indicatrix of scattering, the diffuse reflection coefficient r and the brightness coefficient ϱ, which are related by

$$r = \xi \varrho ,$$

will both depend not only on the angles of incidence and reflection, but also on the difference of the azimuthal angles of the reflected and incident rays. Thus

$$\varrho = \varrho(\eta, \xi; \phi - \phi_0) . \tag{33.31}$$

The application of the principle of invariance in this general case is made in the same way as in the case of a spherical indicatrix of scattering. The formulae in the general case are merely more complex. The transformations made are, however, the same in principle. We shall therefore not give these transformations here, but only the final results.

The brightness coefficient ϱ is represented as a sum:

$$\varrho(\eta, \xi; \phi - \phi_0) = \sum_{m=0}^{n} f_m(\eta, \xi) \cos m(\phi - \phi_0) ,$$

where the number n is equal to the degree of the highest Legendre polynomial, in the expansion (33.6) of the indicatrix of scattering, whose coefficient is not zero.

The functions $f_m(\eta, \xi)$ are expressed in terms of some auxiliary functions $\phi_{i,m}(\eta)$ $(i = m, m+1, \ldots, n)$, each of which depends on only one variable. These auxiliary functions are determined from systems of functional equations. For each function with a given m we have a separate system of functional equations which is solved independently of the others.

The various $f_m(\eta, \xi)$ have, in fact, the following structure:

$$f_0(\eta, \xi) = \tfrac{1}{4}\lambda \frac{\phi_{0,0}(\eta)\,\phi_{0,0}(\xi) - x_1\,\phi_{1,0}(\eta)\,\phi_{1,0}(\xi) + \cdots + (-1)^n x_n \phi_{n,0}(\eta)\phi_{n,0}(\xi)}{\eta + \xi},$$

$$\ldots \qquad \ldots$$

$$f_m(\eta, \xi) = \tfrac{1}{4}\lambda \sum_{i=m}^{n} (-1)^{i+m} \frac{\phi_{i,m}(\eta)\,\phi_{i,m}(\xi)}{\eta + \xi}\, x_i \frac{(i-m)!}{(i+m)!},$$

$$\ldots$$

$$f_n(\eta, \xi) = \tfrac{1}{4}\lambda \frac{x_n}{(2n)!} \frac{\phi_{n,n}(\eta)\,\phi_{n,n}(\xi)}{\eta + \xi} .$$

The auxiliary functions appearing in these expressions are determined from the following systems of functional equations:

$$\phi_{0,0}(\eta) = 1 + \tfrac{1}{2}\lambda\,\eta\,\phi_{0,0}(\eta)\int_0^1 \frac{\phi_{0,0}(\xi)}{\eta+\xi}\,d\xi - \tfrac{1}{2}\lambda\,x_1\,\eta\,\phi_{1,0}(\eta)\int_0^1 \frac{\phi_{1,0}(\xi)}{\eta+\xi}\,d\xi + \ldots + (-1)^n x_n\,\eta\,\phi_{n,0}(\eta)\int_0^1 \frac{\phi_{n,0}(\xi)}{\eta+\xi}\,d\xi,$$

$$\phi_{1,0}(\eta) = \eta - \tfrac{1}{2}\lambda\,\eta\,\phi_{0,0}(\eta)\int_0^1 \frac{\phi_{0,0}(\xi)\,\xi}{\eta+\xi}\,d\xi + \tfrac{1}{2}\lambda\,x_1\,\eta\,\phi_{1,0}(\eta)\int_0^1 \frac{\phi_{1,0}(\xi)\,\xi}{\eta+\xi}\,d\xi + \ldots + (-1)^{n-1} x_n\,\eta\,\phi_{n,0}(\eta)\int_0^1 \frac{\phi_{n,0}(\xi)\,\xi}{\eta+\xi}\,d\xi,$$

$$\ldots \qquad \ldots$$

$$\phi_{n,0}(\eta) = P_n(\eta) + (-1)^n\,\tfrac{1}{2}\lambda\,\eta\,\phi_{0,0}(\eta)\int_0^1 \frac{\phi_{0,0}(\xi)\,P_n(\xi)}{\eta+\xi}\,d\xi + (-1)^{n+1} x_1\,\eta\,\phi_{1,0}(\eta)\int_0^1 \frac{\phi_{1,0}(\xi)\,P_n(\xi)}{\eta+\xi}\,d\xi + \ldots + x_n\,\eta\,\phi_{n,0}(\eta)\int_0^1 \frac{\phi_{n,0}(\xi)\,P_n(\xi)}{\eta+\xi}\,d\xi\,.$$

Similar systems exist for the functions with other values of the second suffix m; instead of the Legendre polynomials we have the associated Legendre polynomials $P_n^m(\eta)$. In particular, $\phi_{n,n}(\eta)$ satisfies the equation

$$\phi_{n,n}(\eta) = (1-\eta^2)^{n/2} + \tfrac{1}{2}\lambda\,\frac{x_n}{(2n)!}\,\eta\,\phi_{n,n}(\eta)\int_0^1 \frac{\phi_{n,n}(\xi)\,(1-\xi^2)^{n/2}}{\eta+\xi}\,d\xi\,.$$

Although these systems of functional equations at first sight appear very complicated, it is easy to solve them numerically by the method of successive approximations, and in any particular case they give comparatively simple results. Thus, in the case of the simplest non-spherical indicatrix of scattering,

$$x(\cos\gamma) = 1 + x_1\cos\gamma\,,$$

the brightness coefficient is

$$\varrho(\eta, \xi; \phi - \phi_0) = \tfrac{1}{4}\lambda\, \frac{\phi_{0,0}(\eta)\,\phi_{0,0}(\xi) - x_1\,\phi_{1,0}(\eta)\,\phi_{1,0}(\xi)}{\eta + \xi} + $$

$$+ \tfrac{1}{4}\lambda\, x_1\, \frac{\phi_1{}'(\eta)\,\phi_1{}'(\xi)}{\eta + \xi}\cos(\phi - \phi_0)\,,$$

where the first two auxiliary functions $\phi_{0,0}(\eta)$ and $\phi_{1,0}(\eta)$ are determined from the system

$$\phi_{0,0}(\eta) = 1 + \tfrac{1}{2}\lambda\,\eta\,\phi_{0,0}(\eta)\int_0^1 \frac{\phi_{0,0}(\xi)}{\eta + \xi}\,\mathrm{d}\xi - \tfrac{1}{2}\lambda x_1\,\eta\,\phi_{1,0}(\eta)\int_0^1 \frac{\phi_{1,0}(\xi)}{\eta + \xi}\,\mathrm{d}\xi\,,$$

$$\phi_{1,0}(\eta) = \eta - \tfrac{1}{2}\lambda\,\eta\,\phi_{0,0}(\eta)\int_0^1 \frac{\phi_{0,0}(\xi)\,\xi}{\eta + \xi}\,\mathrm{d}\xi + \tfrac{1}{2}\lambda x_1\,\eta\,\phi_{1,0}(\eta)\int_0^1 \frac{\phi_{1,0}(\xi)\,\xi}{\eta + \xi}\,\mathrm{d}\xi\,,$$

and the function $\phi_1{}'(\eta)$ is found from the equation

$$\phi_1{}'(\eta) = \sqrt{(1 - \eta)} + \tfrac{1}{4}\lambda x_1\,\eta\,\phi_1{}'(\eta)\int_0^1 \frac{\phi_1{}'(\xi)}{\eta + \xi}\sqrt{(1 - \xi^2)}\,\mathrm{d}\xi\,.$$

These equations are easily solved, and the auxiliary functions tabulated by means of the numerical solution enable us to calculate the brightness coefficient at once.

We may add that the theory shows that not only the diffuse reflection coefficient of the atmosphere but also the diffuse transmission coefficient for a very large optical thickness of the atmosphere are determined in terms of the same auxiliary functions.

The systems of functional equations given here for the auxiliary functions $\phi_{i,m}(\eta)$ are systems of *non-linear* equations. However, it has been shown in the work of V. V. Sobolev that these auxiliary functions satisfy certain *linear* integral equations [157].

The calculations given above do not take into account the phenomenon of polarisation of light in individual scattering processes or the consequent polarisation of the diffusely reflected and diffusely transmitted radiation. The exact theory of the scattering of light in a foggy atmosphere, taking polarisation into account, was developed by V. V. Sobolev in 1943. The results of his investigation were later published [156].

Chapter 34. Comparison of theory and observation

1. **The albedo of a planetary atmosphere.** If we know the intensity of the diffuse radiation emergent in every direction from a planetary atmosphere, we can calculate the total flux of diffusely reflected light from 1 cm^2 of the boundary of the atmosphere. This is clearly

$$\pi H = \int_0^1 \mathrm{d}\eta \int_0^{2\pi} I(\eta, \xi, \phi - \phi_0)\, \eta\, \mathrm{d}\phi\,.$$

In the case of a spherical indicatrix of scattering, I does not depend on the azimuthal angle, and hence

$$\pi H = 2\,\pi \int_0^1 I(\eta,\xi)\, \eta\, \mathrm{d}\eta = 2\,\pi\, S \int_0^1 r(\eta,\xi)\, \eta\, \mathrm{d}\eta\,.$$

Introducing the brightness coefficient ϱ instead of the diffuse reflection coefficient r, we have

$$\pi H = 2\,\pi\, S\, \xi \int_0^1 \varrho(\eta,\xi)\, \eta\, \mathrm{d}\eta\,.$$

On the other hand, since the intensity of the direct rays of the Sun is πS, the flux of these rays incident on 1 cm^2 *of the boundary of the atmosphere* is $\pi S \xi$, where ξ is, as before, the cosine of the angle of incidence of the solar rays. Dividing the flux of emergent radiation πH by the flux of incident radiation, we obtain the albedo of the planetary atmosphere:

$$A = \pi H / \pi S\, \xi = 2 \int_0^1 \varrho(\eta,\xi)\, \eta\, \mathrm{d}\eta\,. \tag{34.1}$$

Let us calculate the albedo A, which will in general depend on the angle of incidence, for the case of a spherical indicatrix of scattering. To do this, we substitute (33.27) in (34.1):

$$A = \tfrac{1}{2}\,\lambda \int_0^1 \frac{\phi(\eta)\,\phi(\xi)}{\eta + \xi}\, \eta\, \mathrm{d}\eta\,. \tag{34.2}$$

Putting

$$\frac{\eta}{\eta + \xi} = 1 - \frac{\xi}{\eta + \xi}\,,$$

we find

$$A = \tfrac{1}{2}\,\lambda\, \phi(\xi) \int_0^1 \phi(\eta)\, \mathrm{d}\eta - \tfrac{1}{2}\,\lambda\, \xi \int_0^1 \frac{\phi(\eta)\,\phi(\xi)}{\eta + \xi}\, \mathrm{d}\eta\,. \tag{34.3}$$

The second term on the right-hand side of (34.3) can be found directly from (33.28) by renaming the variables, and is $\phi(\xi) - 1$. Taking this into account, and using the value of $\int_0^1 \phi(\eta)\, d\eta$ which we calculated from (33.28), we obtain

$$A = 1 - \phi(\xi) \sqrt{(1-\lambda)} . \tag{34.4}$$

It is easy to calculate from this formula the numerical values of the albedo for various ξ and λ, remembering of course that $\phi(\xi)$ depends on the parameter λ. Similar calculations can be performed for a non-spherical indicatrix of scattering. In Table 31 we give the numerical values of the albedo for the case of a spherical indicatrix of scattering.

Table 31

λ / ξ	0·4	0·5	0·6	0·7	0·8	0·9	1·0
0·0	0·23	0·29	0·37	0·45	0·55	0·68	1·00
0·1	0·18	0·24	0·31	0·39	0·49	0·63	1·00
0·2	0·16	0·21	0·28	0·35	0·45	0·59	1·00
0·3	0·14	0·19	0·25	0·32	0·42	0·56	1·00
0·4	0·13	0·17	0·23	0·30	0·39	0·53	1·00
0·5	0·12	0·16	0·21	0·28	0·37	0·51	1·00
0·6	0·11	0·15	0·20	0·26	0·35	0·49	1·00
0·7	0·10	0·14	0·19	0·25	0·33	0·47	1·00
0·8	0·10	0·13	0·17	0·23	0·31	0·45	1·00
0·9	0·09	0·12	0·16	0·22	0·30	0·43	1·00
1·0	0·08	0·12	0·15	0·21	0·29	0·41	1·00

If we know the albedo A for light incident on a given part of the planetary atmosphere at a given angle, we can easily calculate what is called the *spherical albedo* also. By the spherical albedo is meant the ratio of the amount of radiation reflected by the whole planet (from its atmosphere) to the amount incident on the planet (at the boundary of its atmosphere).

The total amount of radiation incident on the planet is clearly $\pi R^2 \cdot \pi S$, where R is the planet's radius. The diffusely reflected energy is given by the integral

$$2\,\pi R^2 \int_0^R \pi H(r)\, d \sqrt{[1-(r/R)^2]} ,$$

where r is the projected distance from the centre of the illuminated disc to some point on the planet's surface. But

$$r/R = \sin \theta_0 = \sqrt{(1-\xi^2)} ,$$

and

$$\pi H(r) = A(\xi)\, \xi\, \pi S\,.$$

Hence the total amount of reflected energy is

$$2\pi R^2 \pi S \int_0^1 A(\xi)\, \xi\, d\xi\,,$$

whence we have for the spherical albedo (denoting it by A^*)

$$A^* = 2 \int_0^1 A(\xi)\, \xi\, d\xi\,. \tag{34.5}$$

In the case of a spherical indicatrix of scattering, we thus have, from (34.4),

$$A^* = 1 - 2\sqrt{(1-\lambda)} \int_0^1 \phi(\xi)\, \xi\, d\xi\,. \tag{34.6}$$

The values of the spherical albedo A^* calculated from this formula, are given in Table 32.

Table 32

λ	0·4	0·5	0·6	0·7	0·8	0·9	1·0
A^*	0·11	0·15	0·19	0·26	0·34	0·48	1·0

Thus we can determine the spherical albedo from λ and, conversely, if the indicatrix of scattering is known, λ can be found from A^*.

2. Comparison with observation. A comparison of the theoretical calculations, given above for a spherical indicatrix of scattering, with the results of absolute determinations of the diffuse reflection coefficient for Jupiter made by V. V. Sharonov, has shown that the observed brightness distribution over Jupiter's disc is in good agreement with the theory for $\lambda = 0{\cdot}969$.

However, it must be remarked that, within the limits of observational accuracy, the same brightness distribution over Jupiter's disc can be explained by assuming any other, non-spherical, indicatrix of scattering, though for some other value of λ. This example shows that, for the planets whose phase angle hardly varies (Jupiter, Saturn, etc), it is difficult to determine the indicatrix of scattering from single observations of the brightness distribution.

The situation is quite different for Venus, which we observe at all possible phase angles. In this case the form of the indicatrix of scattering has a fairly important effect even on the light curve of the planet. V. V. Sobolev, using the observed light curve constructed by

H. N. Russell, was able to find the indicatrix of scattering for the atmosphere of Venus by comparison with the approximate theory of diffuse reflection which he had devised for an arbitrary indicatrix of scattering. It was found that the parameter λ for the atmosphere of this planet is 0·989, and the indicatrix of scattering is a function whose numerical values are given in Table 33.

Table 33

γ	0°	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	165°	180°
$x(\cos\gamma)$	7·2	6·0	2·7	1·5	1·0	0·07	0·51	0·40	0·35	0·41	0·56	0·65	0·70

It is interesting to compare the indicatrix of scattering in the atmosphere of Venus with that in the Earth's atmosphere. The scattering properties of the Earth's atmosphere are estimated not from the processes of diffuse reflection, but by investigating those of diffuse transmission. Here, despite the comparatively small optical thickness of the Earth's atmosphere (when free from cloud), which is of the order of 0·3 in the visible part of the spectrum, it is nevertheless, important, as V. V. Sobolev has shown, to take into account scattering not only of the first order, but also of higher orders, i. e. to use a fairly rigorous theory. The determination of the indicatrix of scattering in the Earth's atmosphere has been carried out by many authors. The most interesting results have been obtained by E. V. Pyaskovskaya-Fesenkova and V. V. Sobolev. The latter made his determination from observations on the brightness distribution over the sky, carried out in the summer of 1943 at Yelabuga by workers from the astronomical observatory of the Leningrad State University. The results of his determination are given in Table 34.

Table 34

γ	0°	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	165°	180°
$x(\cos\gamma)$	4·6	3·3	1·9	1·3	0·94	0·75	0·64	0·65	0·72	0·85	1·03	1·1	1·2

A comparison of Tables 33 and 34 shows that the indicatrix of scattering in the atmosphere of Venus is more elongated in the forward direction (small angles of deviation) than that in the Earth's atmosphere.

It is easy to see that an idea of the degree of elongation of the indicatrix of scattering in the forward direction is given, first of all, by the coefficient x_1 of the polynomial $P_1(\cos\gamma) = \cos\gamma$ in the expansion formula (33.6). For a spherical indicatrix, or for one satisfying the condition $x(\gamma) = x(180° - \gamma)$, i. e. symmetrical about the direction $\gamma = 90°$, we have $x_1 = 0$. For an indicatrix elongated forwards $x_1 > 0$, and for one elongated backwards $x_1 < 0$. For an indicatrix which is

very sharply elongated forwards, like a needle, i. e. when the scattering is confined to small deviations from the original direction, $x_1 = 3$. For a similar elongation backwards, $x_1 = -3$. Thus the value of x_1 gives a first rough conception of the elongation of the indicatrix.

If the scattering in the atmosphere of the Earth or Venus were *purely molecular*, the indicatrix of scattering would have the form, deduced by Lord RAYLEIGH,

$$x(\cos\gamma) = \tfrac{3}{4}\,(1 + \cos^2\gamma)\;.$$

For RAYLEIGH's indicatrix, the coefficient x_1 is zero, i. e. *there is no one-sided elongation.* The considerable elongation of the indicatrices of scattering in the atmospheres of the Earth and Venus indicates that in both cases, besides purely molecular scattering, which of course always occurs, there is additional scattering of light by larger liquid or solid particles suspended in the atmosphere. Since the number of suspended particles in different parts of the atmosphere may not be proportional to the number of scattering molecules of gas, the indicatrix should vary with height in the atmosphere, which greatly complicates the theory. In particular, on the Earth the haze due to fine suspended particles is more concentrated in the lower layers of the atmosphere.

The degree of elongation of the indicatrix of scattering depends on the dimensions of the scattering particles. In particular, transparent particles (drops of water and of other liquids) have an indicatrix of scattering which is the more elongated forwards, the greater their diameter. This is why, on comparing Tables 33 and 34, we can say that the scattering in the atmosphere of Venus is due to larger particles than that in the Earth's atmosphere. Here, however, we must remember that, in deriving the indicatrix of scattering for the Earth's atmosphere, *the case where clouds are absent* was considered. However, fairly large water drops are found in clouds. In the case of the atmosphere of Venus we have continuous cloudiness, and the indicatrix found is characteristic mainly of the properties of the particles in these clouds.

The theory of scattering of light, developed above for the case of a medium of infinite optical thickness, can be generalised to media of finite optical thickness. Here, besides the coefficients of diffuse reflection of light, we have to deal also with those of diffuse transmission. For this reason the number of auxiliary functions of one variable, from which the functions giving the values of these coefficients are constructed, is doubled. The number of equations in the system of functional equations for these auxiliary functions is correspondingly doubled also.

The limiting case of large optical thicknesses occupies a special place. The diffuse reflection coefficient tends, as $\tau \to \infty$, to the coefficient

of diffuse reflection for an infinitely thick layer. The diffuse transmission coefficient tends to the product $f(\tau_0)\, u(\theta)$, where $u(\theta)$ depends only on θ and not on the direction of the original radiation entering the medium from the other side. Thus, in the limit, the relative distribution in direction of the transmitted radiation becomes constant.

We are concerned with such a case of diffuse transmission of radiation through a medium of very great optical thickness when we consider the Earth's atmosphere with uniform and fairly heavy cloud. Observation shows that the relative distribution of brightness over a cloudy sky, when the cloud is continuous and uniform, varies little from day to day, if we omit the effect of such agencies as the covering of the Earth's surface with snow. If the Earth's surface is not snow-covered, the reflection from it can be neglected in rough calculations. Thus the conditions of summer and winter must be distinguished in considering the distribution of radiation in the atmosphere. A certain constancy of the brightness distribution over the sky in each set of conditions shows that certain optical properties of the water particles forming the clouds are constant. These properties must include the nearness of λ to unity. In other words, we are concerned here with almost pure scattering. The fact that λ is close to unity is confirmed in the following ways.

(1) The daytime illumination from a cloudy sky varies comparatively little, despite the fact that the optical thickness of the cloud is sometimes very great. Any noticeable true absorption (i. e. a noticeable deviation of λ from 1) would necessarily cause, at such great optical thicknesses, a very great decrease in the daytime illumination at the Earth's surface, and this is not observed.

(2) The albedo of the cloud is shown by observations from aeroplanes to be high.

(3) In the presence of snow, the distribution of brightness over the sky approaches uniformity. In this case the snow, being a scattering medium, forms, roughly speaking, an extension of the cloud, and the observer is at the boundary of two semi-infinite scattering media. It can be shown that the brightness distribution in various directions will appear isotropic to such an observer if $\lambda = 1$ in both media.

The indicatrix of scattering in clouds is very elongated in the forward direction.

As we have said above, the clouds in our atmosphere consist of a collection of water drops of spherical shape. For drops of spherical shape, the indicatrix of scattering can be calculated theoretically. It depends on the radius of the drop. The greater the ratio a/λ of the radius of the drop to the wavelength of the light, the more elongated is the indicatrix. Since, according to present data, the clouds consist

of droplets the majority of which have radii a few times the wavelength of visible light, the theoretical indicatrices for this particle radius must be used. Laboratory investigations on the determination of the indicatrix of scattering for fog particles have confirmed that it is very elongated, and have resulted in good agreement with the theory.

In an actual cloud and in fog, there are particles of various diameters. For this reason, the indicatrix of scattering in a cloud is some weighted mean of the indicatrices for droplets of various sizes. This mean indicatrix will depend on the droplet dimensions which predominate in a given assembly of clouds.

What is valid for water droplets distributed in the Earth's atmosphere holds also for droplets of other liquids, which probably play a large part in the atmospheres of other planets (methane, ammonia).

It must be noted, however, that, although very elongated, the indicatrix of scattering in most clouds is such that deviations by at least two or three degrees are more probable than smaller deviations, i. e. almost rectilinear propagation. If this were not so, we should usually find the limb of the Sun's disc obscured on observing it through thin layers of cloud. Nevertheless, the limb of the Sun's disc continues to remain sharp, in the majority of cases, when observed through such cloud layers.

3. Absorption bands. The scattering of light, as well as the true absorption accompanying it, which phenomena we have been discussing hitherto, take place in all the visible and photographic parts of the spectrum, and are slowly varying functions of the wavelength. However, in the spectra of the major planets we observe absorption bands corresponding to definite molecules. These are due to *monochromatic absorption* in the lines of the band.

The results of observations of these bands are usually expressed as the path length which a ray must traverse in the gas concerned at atmospheric pressure in order to produce a band (or line) of the same equivalent width. If we observed the passage of the direct rays of the Sun through the atmospheric layers of the outer planets, such data would have an immediate physical significance. We observe, however, only the diffuse reflection of the continuous spectrum of the Sun's radiation from the planet's atmosphere. Hence we must attempt to decide more exactly the physical significance of the "equivalent paths", expressed in kilometres at atmospheric pressure, which are obtained from observations of the spectra of planets.

The answer to this question depends on the interrelation between the factors causing the absorption bands and the scattering in the continuous spectrum.

Here two possibilities must be borne in mind:

(1) The scattering of light may be due to the gases forming the planet's atmosphere. It is then of the same nature as the Rayleigh scattering which brings about the blueness of the sky on the Earth. In this case the medium causing the monochromatic absorption is the same as that causing the scattering in the continuous spectrum, since the monochromatic absorption can be due only to the gases which are mixed with one another in the atmosphere. Assuming that there is complete mixing, we can suppose that the presence of absorption lines merely shows that there are minima in the curve giving the dependence of the parameter λ on the frequency. The table above which gives the dependence of the albedo on λ shows that even a small decrease in the value of λ can lead to a marked diminution in the albedo.

In this case, the interpretation of the observations must consist, first of all, in determining the variation of λ with frequency from spectrophotometric data. If the values of λ for various frequencies have been found (including frequencies inside and outside a line), then, neglecting the ratio k_ν/s_ν, we can find k_ν'/s_ν from the formula

$$\lambda_\nu = s_\nu/(s_\nu + k_\nu + k_\nu') ,$$

where s_ν is the scattering coefficient, k_ν the true absorption coefficient in the continuous spectrum, and k_ν' the additional true absorption coefficient in the line. The value of k_ν', referred to one molecule per cm^2, is a quantity which can be determined from experiment. Knowing the ratio k_ν'/s_ν, we can then find s_ν referred to one molecule per cm^2. The determination of the value of s_ν referred to one molecule of the absorbing gas, on comparison with our considerations on the mechanism of scattering, may help to determine the fraction of molecules which absorb radiation in the given line, out of the total number of molecules of the gaseous mixture forming the atmosphere.

(2) The scattering may be due to liquid particles (drops) or to solid particles (small crystals) forming clouds lying in a cloudy layer bounded in altitude on both sides. Then the Sun's rays, before undergoing diffuse reflection from this cloudy layer, must reach it by passing through the overlying layers of the atmosphere, and they must again traverse these overlying layers before reaching the observer. The formation of an absorption band may occur in these overlying layers. In this case, the equivalent path lengths obtained from observation are actually characteristic of the mass of the gas concerned which lies above the cloudy layer.

PART IX.
INTERSTELLAR MATTER

Chapter 35. The dust component of interstellar matter

The investigation of the interstellar matter which causes the absorption of the light of distant stars is linked with the name of the celebrated Russian astronomer Vasiliĭ Yakovlevich Struve, who, in the first half of the nineteenth century, made an attempt to determine the amount of this absorption. In the twentieth century, the problem of the interstellar absorption of light has attracted the attention of several Russian astronomers, and has been investigated with particular vigour during the last thirty years. Many of the important ideas which have arisen in this branch of science in recent years are due to Soviet astronomers.

The existence of interstellar matter in our stellar system and in the other galaxies can be seen from a large number of facts. We shall discuss some of these below.

The discovery of the spectral line of hydrogen with the radio wavelength of 21 cm has been of very great importance in the study of interstellar matter. The possibility of observing this line was predicted by H. C. van de Hulst, and it was independently discovered by C. A. Muller in Holland and by H. I. Ewen and E. M. Purcell in Australia. Numerous observations of the interstellar radiation in this line that have been made recently in Holland and the United States have considerably extended our knowledge of the distribution of interstellar matter.

1. Some basic facts. The following simple fact indicates the existence of matter which absorbs light in the interstellar spaces of our Galaxy. The number of other galaxies (extra-galactic nebulae) brighter than some magnitude m found in one square degree has its greatest value in the direction of the galactic poles. As we approach the galactic equator, this number N_m decreases, and in the directions lying in the zone of the Milky Way no extra-galactic nebulae at all are observed, as a rule. Only in some sections of the Milky Way, called "galactic windows", are a small number of other galaxies observed. Thus the zone of the Milky Way is a "zone of avoidance" for the extra-galactic nebulae.

Since the other galaxies are systems independent of our Galaxy, this diminution in the number of other galaxies as we approach the galactic equator must be regarded as a spurious phenomenon, caused by the presence of a layer of absorbing matter near the plane of symmetry of the Galaxy. We can suppose that the extent of the layer of "dark" absorbing matter in the direction perpendicular to the galactic plane is very small in comparison with its extent in directions along this plane, i. e. that the dark matter is concentrated about the plane of symmetry of the Galaxy, as are giant stars and ordinary dwarfs, and that it consists of plane-parallel layers, i. e. that the density of the absorbing matter is a function only of the distance z from the plane of symmetry of the Galaxy. It is then not difficult to explain the observed variation in the number N_m of extra-galactic nebulae with galactic latitude.

Under these conditions, if we denote by τ the optical thickness of the galactic absorbing layer for the radiation of those extra-galactic nebulae which lie at the galactic pole, then the light of the extra-galactic nebulae observed in galactic latitude b will pass through an optical thickness τ cosec b, and consequently will be attenuated by a factor $e^{-\tau \operatorname{cosec} b}$. If we determine, for various galactic latitudes b, the number N_m of extra-galactic nebulae down to a given *apparent* magnitude m, there will correspond to our counts at various latitudes various corrected limiting stellar magnitudes,

$$m_0 = m + 2{\cdot}5 \log_{10} e^{-\tau \operatorname{cosec} b}$$

or

$$m_0 = m - 2{\cdot}5\, M\, \tau \operatorname{cosec} b = m - 1{\cdot}09\, \tau \operatorname{cosec} b\,, \qquad (35.1)$$

where M is the modulus of natural logarithms.

On the other hand, counts of extra-galactic nebulae down to various apparent magnitudes in a fixed direction show that N_m increases with m approximately according to the law

$$N_m = N_1 \times 10^{0{\cdot}6m}\,, \qquad (35.2)$$

where N_1 is a constant for the given region of the sky; this relation holds at least as far as 20^{m}.

The relation (35.2) means that the number of extra-galactic nebulae N_B down to some value of the total luminosity B is *inversely proportional to* $B^{3/2}$ (by the definition of stellar magnitude, the total luminosity is proportional to $10^{-0{\cdot}4m}$). This, however, can only be regarded as an indication of a more or less uniform large-scale distribution of the other galaxies in extra-galactic space.

If the spatial distribution of the extra-galactic nebulae is uniform or nearly so, we must have for the stellar magnitudes, corrected on account of absorption, a distribution

$$N_m = N_0 \times 10^{0 \cdot 6 m_0}, \tag{35.3}$$

where N_0 is a constant which is the same for the whole sky. Substituting the value of m_0 from (35.1) in (35.3), we find

$$N_m = N_0 \times 10^{-0 \cdot 65 \tau \operatorname{cosec} b} \times 10^{0 \cdot 6 m}. \tag{35.4}$$

Comparing (35.4) and (35.2), we see that

$$N_1 = N_0 \times 10^{-0 \cdot 65 \tau \operatorname{cosec} b}. \tag{35.5}$$

Thus, if we plot on the axes of co-ordinates $\log_{10} N_1$ and cosec b, the relation obtained between these quantities should be a straight line, the slope of which should be 0·65 τ. The counts of extra-galactic nebulae, on being reduced in this way, therefore make it possible to determine the optical thickness traversed by the radiation from the extra-galactic nebulae observed at the pole of the Galaxy. It is evident that the value obtained (assuming that the Earth lies close to the plane of symmetry of the Galaxy) will be equal to half the total optical thickness of the whole absorbing layer. On multiplying τ by 1·09 we have the same half-thickness expressed in stellar magnitudes.

A determination carried out by E. P. Hubble has given $t = 0^{\mathrm{m}}{\cdot}25$ for this half-thickness. However, an investigation by P. P. Parenago, based on a detailed analysis of counts of extra-galactic nebulae, has led to a more reliable value, $t = 0^{\mathrm{m}}{\cdot}34$. The latter estimate shows, in particular, that at a galactic latitude $b = 10°$ the absorption should reach two magnitudes, and this means that at this latitude the average number of extra-galactic nebulae of a given magnitude in one square degree should be sixteen times less than at the poles.

When $|b|$ becomes less than five degrees, i.e. when the ray lies almost in the plane of the Galaxy, the optical thickness of the layer becomes so great that we see hardly any extra-galactic nebulae in such low latitudes.

For the plane of the Galaxy itself, this conception of the absorbing layer allows an interesting conclusion to be drawn concerning the surface brightness of the Milky Way. Assuming that, as well as the layers of absorbing matter, those of radiating matter, i.e. the distribution of stars, are plane-parallel, we can say that the plane of symmetry of the Galaxy ($z = 0$) is characterised by a constant macroscopic absorption coefficient α and a constant macroscopic emission coefficient j. That is,

$4\pi j$ is the amount of energy emitted by all the stars in unit volume in this plane, in unit time and in all directions. Then we have as usual for the intensity of a ray coming to us in the plane of the Galaxy

$$I = \int_0^\infty e^{-\tau} j \, \mathrm{d}r , \tag{35.6}$$

where $\mathrm{d}r$ is an element of the path of the ray and τ is the optical thickness of the absorbing layer measured from the observer to r. For τ thus introduced we can write

$$\tau = \alpha r , \tag{35.7}$$

where α is the mean volume absorption coefficient. Putting

$$j \, \mathrm{d}r = (j/\alpha) \, \mathrm{d}\tau , \tag{35.8}$$

we have instead of (35.6)

$$I = (j/\alpha) \int_0^\infty e^{-\tau} \, \mathrm{d}\tau = j/\alpha . \tag{35.9}$$

We have extended the integration in (35.9) from 0 to ∞, since the optical thickness of the absorbing layer in the directions where $b = 0°$ is so great that the replacement of the finite upper limit by infinity is unimportant. The value of j can be found by adding the luminosities of the stars in some volume, and then dividing the sum by this volume and by 4π. For the volume concerned we can take a sphere centred at the Sun, of radius 5 parsecs. If there are any faint stars inside this sphere which are still unknown to us, they cannot cause a noticeable error in the calculation of the sum of the luminosities. Thus j can be reliably determined. On the other hand I, that is the mean brightness of the stellar component of the Milky Way, can be determined from observation. We can hence find the mean value of α, which, as calculation shows, is approximately one stellar magnitude per kiloparsec.

We notice that, by introducing the macroscopic emission coefficient j for any stellar system, we can generally establish an important relation concerning the surface brightnesses of the galaxies, which holds in cases where the optical thickness of the system is small in the given direction, and cannot be equated to infinity:

$$I = \int_0^{\tau_0} e^{-\tau} (j/\alpha) \, \mathrm{d}\tau = P(1 - e^{-\tau_0}) , \tag{35.10}$$

where τ_0 is the optical thickness of the system in the direction concerned. Here P is some mean value of j/α over the path of the ray in the system.

It is interesting that the central parts of the elliptical galaxies, like the central parts of the nuclei of the spiral galaxies, have a surface brightness (intensity) which in many cases exceeds by a factor of a hundred the observed mean surface brightness of the Milky Way. This means that the ratio j/α in them is, on the average, a hundred times greater than it is in the part of the Galaxy containing the Sun and the nearest stars. In other words, in the elliptical nebulae and in the nuclei of the spiral nebulae, the part played by the absorbing matter is negligible in comparison with the exterior parts of our Galaxy (where the Sun lies).

As we shall see below, the distribution of absorbing matter in the part of the Galaxy where the Sun is, and in the exterior parts of the other spiral galaxies, and also in galaxies of irregular shape, is not only non-uniform, but cannot even be supposed continuous. In other words, the volume absorption coefficient α changes discontinuously. This means that the interstellar matter has *a patchy, cloudy structure*, and if we can speak of an absorbing layer in the Galaxy it is *a layer of absorbing clouds*.

In the light of these considerations it is easy to understand why, in cases where we view a spiral nebula "edge-on", i.e. when we are ourselves close to the principal plane of symmetry of the system, we see that its bright nucleus is crossed by a dark band. The reason is that the exterior parts (including the spiral arms), for which the mean value $P = j/\alpha$ is of the same order as in the neighbourhood of the Sun, are in this case projected on the nucleus, which has a high surface brightness. The surface brightness in these parts of the projection, which is due mainly to the light of the exterior stars (the light of the nucleus is greatly attenuated, in consequence of the large optical thickness of the absorbing layer), should be of the same order as the surface brightness of the Milky Way, i.e. less than that of the nucleus by a factor of the order of ten. Thus, the fact that such dark bands are observed superposed on the bright nuclei of nebulae seen "edge-on" needs no special hypothesis that, say, absorbing matter is present around the galaxy in question, but only the assumption that the conditions in the outer parts of this galaxy are similar to those which exist in the exterior parts of our own Galaxy. For the same reason, the order of magnitude of the brightness of the Milky Way in the direction of the nucleus of the Galaxy (in the constellation of Sagittarius) is the same as in other directions in the Milky Way.

Another very important fact related to the interstellar matter in the Galaxy is the reddening of the light of distant stars. This phenomenon is due to the fact that the coefficient of interstellar absorption always increases as the wavelength decreases, and this brings about a reddening of the stars compared with the normal colour for the spectral class concerned. Since the light penetrates through greater optical thicknesses

in traversing long segments of its path close to the plane of symmetry of the Galaxy, it is natural to expect the greatest reddening for stars observed at very great distances and at low galactic latitudes. Among such stars, those of classes O and B and the long-period cepheids are suitable from the point of view of the possibility of determining their colours.

The colour indices of stars of classes O and B have been studied by J. STEBBINS, C. M. HUFFER and A. E. WHITFORD, and the colour indices of a large number of long-period cepheids have been determined by G. S. BADAL'YAN, and independently by M. A. VASHAKIDZE. From results on the colour indices of nearer stars of these types, the normal colour indices $(C_i)_0$ (i.e. the colour indices free from the effects of absorption), have been determined for each sub-class from O to B 0 (and for cepheids). By subtracting these normal colour indices from the observed ones, the colour excesses C_e were obtained:

$$C_e = C_i - (C_i)_0 \,. \tag{35.11}$$

The appearance of the colour excesses is the result of the unequal absorption of photographic and visible light. If the absorption, in the photographic range, of the light of a given star, expressed in stellar magnitudes, is denoted by A_p, and that in the visible range, expressed in the same way, by A_v, then

$$C_e = A_p - A_v \,, \tag{35.12}$$

provided that the colour indices themselves, from which the colour excesses are determined, have been found by comparing the photographic and visual magnitudes.

On the other hand, the absorption A_p can also be expressed in terms of the optical thickness traversed by the star's radiation in the photographic wavelengths:

$$A_p = 1{\cdot}09\, \tau_p \,, \tag{35.13}$$

since both A_p and τ_p are characteristic of the absorption. For the transmission of the thickness traversed is expressed in terms of these quantities by

$$(2{\cdot}512)^{-A_p} = e^{-\tau_p} \,. \tag{35.14}$$

Similarly

$$A_v = 1{\cdot}09\, \tau_v \,, \tag{35.15}$$

where τ_v is the optical thickness traversed by the light of the star in the visible frequencies.

Hence we can write instead of (35.12)

$$C_e = 1{\cdot}09\,(\tau_p - \tau_v)\,. \tag{35.16}$$

On the other hand, introducing the absorption coefficients α_p and α_v for the two regions of the spectrum, we can write

$$\tau_p - \tau_v = \int_0^r (\alpha_p - \alpha_v)\,\mathrm{d}r \tag{35.17}$$

and

$$\tau_p = \int_0^r \alpha_p\,\mathrm{d}r\,, \tag{35.18}$$

whence, by (35.15) and (35.16),

$$C_e = 1{\cdot}09 \int_0^r (\alpha_p - \alpha_v)\,\mathrm{d}r\,, \tag{35.19}$$

$$A_p = 1{\cdot}09 \int_0^r \alpha_p\,\mathrm{d}r\,. \tag{35.20}$$

It is very unfortunate that we do not at present have any method, other than statistical ones, of directly determining A_p. Hence a direct determination of A_p for individual stars is extremely difficult. For A_p appears in the equation relating the apparent and absolute photographic magnitudes:

$$m = M + 5\log_{10} r - 5 + A_p\,, \tag{35.21}$$

and to find A_p we must be able to determine independently the absolute magnitude and the distance. For instance, the expression (35.21) could be used in cases where the parallax has been determined trigonometrically and the absolute magnitude spectroscopically. Unfortunately, however, the trigonometric parallaxes have not been determined at all accurately for distances over 100 parsecs. The absorption at smaller distances is very small. Hence another possibility is widely used, namely to calculate A_p from the colour excess, which is obtained, as we have said, from observation.

Thus, provided that the ratio of the absorption coefficients $\mu = \alpha_v/\alpha_p$ in the two regions of the spectrum is constant in space, we have by (35.19)

$$C_e = 1{\cdot}09\,(1-\mu) \int_0^r \alpha_p\,\mathrm{d}r\,, \tag{35.22}$$

whence, on comparing with (35.20), we find

$$C_e = (1-\mu)\, A_p. \tag{35.23}$$

Thus, in this case, the total absorption in photographic light for each star would be obtained from its colour excess by multiplying by the same constant

$$\gamma = 1/(1-\mu)\,. \tag{35.24}$$

This method is the most frequently applied at present to determine the total absorption of the photographic light of a given star. The questions of the physical conditions under which the ratio μ can be constant, and whether these conditions are in fact fulfilled, are very closely connected with the problem of the nature of the matter which absorbs light.

As we shall point out below, the interstellar matter produces some absorption lines, belonging to the atoms and ions of certain elements, in the spectra of individual stars lying close to the galactic plane. Thus the presence in the composition of the interstellar matter of gaseous masses producing these lines is indubitable. However, the question arises whether matter in the gaseous state can cause both *selective* and *general absorption of the whole continuous spectrum* such as we observe, i.e. more intense in the region of short wavelengths and less so in the region of long wavelengths.

To answer this question, we must decide what is the cause of the opacity of the interstellar gas in the continuous spectrum. As we have seen in the early Parts of this book, there are three main sources of opacity of a gas. One of these is the bound-free transitions of the electrons, another is the free-free transitions, and the third is the scattering of light by free electrons. The first two of these, which depend in a complicated manner on the wavelength, for a given temperature, give a mass absorption coefficient which is proportional to the electron pressure. The third source gives, for the almost total ionisation which must exist in the interstellar gas, a mass scattering coefficient which is practically constant. Hence, when the total density, and therefore the electron pressure, decrease, the part played by the first two sources of opacity diminishes markedly in comparison with that of the third source. Even in the solar corona, the radiation in the continuous spectrum is entirely due to the scattering of light by free electrons, and the bound-free and free-free transitions of electrons do not produce any noticeable absorption (with subsequent re-emission) of the Sun's radiation by the corona in the visible and photographic regions of the spectrum.

This is even more true of the interstellar gas, which must have a density many millions of times less than the solar corona. This means that the opacity of the interstellar gas in the continuous spectrum must be entirely

due to the scattering of light by free electrons. This scattering, however, is independent of the wavelength, and hence it cannot explain the interstellar absorption, which, as we have seen, is selective. Thus we can assert that the existence of the interstellar gas cannot account for the observed general absorption of starlight in the continuous spectrum. The question may arise whether the scattering of light by free electrons cannot explain at least a part of the observed attenuation of the light of the stars, while the remainder is due to some other very selective agency. In order to answer this question, let us first attempt to estimate the mass of gas necessary to bring about the attenuation of the light of the stars that is observed in photographic wavelengths, assuming that the attenuation is due to this scattering of light by free electrons. We can suppose that the mass of interstellar gas in the entire Galaxy must be equal to the cross-section πR^2 of the Galaxy in its plane of symmetry (where R is the radius of the Galaxy) multiplied by the mean mass in a column of cross-section 1 cm^2 perpendicular to the galactic plane, i.e.

$$\mathsf{M}_g = \pi R^2 \int_{-\infty}^{\infty} \varrho \, \mathrm{d}z ,$$

where the integration is over the z co-ordinate, which is measured perpendicular to the galactic plane. Introducing the scattering coefficient s referred to unit mass, we obtain

$$\mathsf{M}_g = \pi R^2 \frac{1}{s} \int_{-\infty}^{\infty} s \varrho \, \mathrm{d}z = 2 \pi R^2 \tau_z / s , \tag{35.25}$$

where τ_z is the optical half-thickness of the Galaxy. Here we have supposed s to be constant. We can take hydrogen as the principal constituent of the interstellar gas; the overwhelming majority of the hydrogen atoms are ionised. Hence

$$s = n_{\mathrm{H}} s_0 ,$$

where s_0 is the scattering coefficient when there is only one free electron in unit volume, and n_{H} is the number of hydrogen atoms (ions in this case) in one gram. Both these quantities are not only constants, but universal constants:

$$n_{\mathrm{H}} = 1/m_{\mathrm{H}} , \quad s_0 = (8 \pi/3) (e^2/mc^2)^2 = 0{\cdot}7 \times 10^{-24} \, \mathrm{cm}^2 .$$

The quantity s can therefore be regarded as constant also. The electrons detached from atoms of other elements do not add much to the value of s thus calculated. Thus, instead of (35.25), we can write

$$\mathsf{M}_g = m_{\mathrm{H}} \, . \, 2 \, \tau_z \, \pi \, R^2 / s_0 \, . \tag{35.26}$$

Taking $R = 10{,}000$ parsecs $= 3 \times 10^{22}$ cm and $2\,\tau_z = 0{\cdot}7$, we obtain, for the mass of absorbing gas in the Galaxy, a value of the order of 10^{45} g, which is several times the total mass of the Galaxy itself. Hence, even without taking into account the observed selective attenuation of light, we can say that free electrons cannot be responsible for even an important part of the optical thickness of the absorbing layer in photographic wavelengths.

Thus we must completely abandon the idea that the gaseous component of the interstellar matter can play any noticeable part in the general observed phenomenon of the attenuation of the radiation of the stars over the whole spectrum.

Since, however, the independent existence of liquid particles in interstellar space is ruled out, because the temperature of a solid or liquid particle in interstellar space, far from any star, must be extremely low (less than 10° K), the conclusion can only be drawn that *the interstellar absorption is caused by solid particles.* Here it must be borne in mind that the chief part in the absorption of light cannot be ascribed to large particles, exceeding (say) 1 mm in linear dimensions. In fact, just as in the hypothesis that free electrons play a part in this phenomenon, the assumption that the dimensions of the particles are so large leads to a very large value for the mass of absorbing matter in the Galaxy, which is inadmissible in galactic dynamics. Moreover, in this case the absorption would be not selective but neutral: when the dimensions of the particles are many times the wavelength, their effect appears as a *screening* of the light of the stars. For in this case geometrical optics is valid, and particles of this kind projected on a star's disc will eclipse separate, though of course very small, parts of the disc. When the number of particles is sufficiently large, the light of the star will be noticeably attenuated, and this will occur, in all wavelengths which are small compared with the dimensions of the particles, in the same proportion, namely the ratio of the part of the disc covered by the particles to the whole disc. In fact, however, we observe *selective* absorption. This can be caused only by particles whose linear dimensions are small in comparison with the wavelength of the light, or at most of the order of the wavelength of visible light.

The physical theory of the scattering of light by such small particles was developed by G. Mie. It was found that not only the dimensions of the particles, but also their physical properties (e.g. their electrical conductivity), are of importance.

If we were able to form any definite idea about the statistics of the dimensions and other properties of the particles of interstellar matter, or, as one generally says, of the *cosmic dust*, then, on the basis of Mie's theory, we could calculate the absorption coefficient as a function

of wavelength. However, since such an idea would be purely conjectural, it is necessary to determine the absorption coefficient as a function of wavelength empirically, and only then to form an idea of the composition of the cosmic dust by attempting to explain the observed dependence of the absorption coefficient on the wavelength.

The determination of the absorption coefficient as a function of wavelength over the spectrum is in principle a simple problem. We have to take two stars with the same physical properties that can be ascertained from their spectral lines. They must not only belong to the same spectral sub-class, but have the same luminosity, and it is therefore necessary that the actual form of the spectral lines (their width) should be the same. When stars similar in their physical properties have been carefully selected in this way, we may expect that not only are their colour temperatures the same, but the energy distribution in the continuous spectrum is also exactly the same. If one of these stars is distant and its light undergoes considerable absorption, the difference in the apparent relative energy distribution in their spectra may be regarded as being due entirely to the selective interstellar absorption.

Let $I_0(\lambda)$ be the relative energy distribution in the spectrum of the near star (free from absorption), and $I_1(\lambda)$ the observed distribution in the spectrum of the more distant star. If there were no absorption, we should have

$$I_1(\lambda) = C\, I_0(\lambda)\,, \tag{35.27}$$

where C is a constant depending on the ratio of the distances of the stars concerned. However, since the light has undergone absorption in the second case, we have, instead of (35.27),

$$I_1(\lambda) = C\, e^{-\tau_\lambda}\, I_0(\lambda)\,, \tag{35.28}$$

where τ_λ is the optical thickness of the path traversed by the light in wavelength λ. From (35.28) we have

$$\tau_\lambda = -\log_e\,[I_1(\lambda)/I_0(\lambda)] + \log_e C\,. \tag{35.29}$$

For τ_λ, however, we have

$$\tau_\lambda = \int\limits_0^r \alpha_\lambda\,\mathrm{d}r = \frac{\alpha_\lambda}{\alpha_{\lambda_1}}\int\limits_0^r \alpha_{\lambda_1}\,\mathrm{d}r = \frac{\alpha_\lambda}{\alpha_{\lambda_1}}\,\tau_{\lambda_1}\,, \tag{35.30}$$

where λ_1 is some chosen wavelength, and it is assumed that the ratio $\alpha_\lambda/\alpha_{\lambda_1}$ is constant, i. e. independent of distance. Introducing (35.30) in (35.29), we find

$$\frac{\alpha_\lambda}{\alpha_{\lambda_1}} = -\frac{1}{\tau_{\lambda_1}}\log_e\frac{I_1(\lambda)}{I_0(\lambda)} + \frac{1}{\tau_{\lambda_1}}\log_e C\,.$$

Since the constant C is usually undetermined, we can say that the ratio $\alpha(\lambda)/\alpha(\lambda_1)$ is determined from observation apart from an additive constant.

The use of observational data has shown that, in the majority of cases, $\alpha(\lambda)/\alpha(\lambda_1)$, and therefore $\alpha(\lambda)$ itself, is a linear function of λ^{-1}, i. e.

$$\alpha(\lambda) = c_1\, \lambda^{-1} + c_2\,, \tag{35.31}$$

where c_2 is, so to speak, the neutral component of the absorption curve. MIE's theory leads to the conclusion that such a dependence of the absorption coefficient on the wavelength in the photographic part of the spectrum is given by metallic particles (the calculations were made for iron and nickel) with diameters of the order of 10^{-5} cm, i. e. less than the wavelength. It is clear that the interstellar dust is composed of particles of very different diameters. The statistics of these diameters ought to be investigated more fully, but the result obtained shows that diameters of the order of 10^{-5} cm are the most effective in interstellar absorption.

It may be mentioned that some authors have attempted, starting from the assumption that $\alpha(\lambda)$ has the general form

$$\alpha(\lambda) = c_1\, \lambda^{-k} + c_2\,, \tag{35.32}$$

to determine the value of k which best agrees with the spectrophotometric results. However, the values of k thus determined do not deviate greatly from unity. O. A. MEL'NIKOV has shown that it is better to take k as varying with the wavelength, since the same particles give various exponents k for different wavelength ranges. Thus, when the wavelength is several times less than the diameter, each particle absorbs neutrally, i. e. $k = 0$. Further investigations are necessary to derive the variation of k with the wavelength range. Only the establishment of the change of k with wavelength makes it possible to draw some preliminary conclusions on the statistics of the dimensions of the cosmic dust particles.

At present, formula (35.31) is generally used in practice. However, the value of the constant c_2 (or, more precisely, of the ratio c_2/c_1) cannot be determined from spectrophotometric observations alone. To determine it, we must have at the same time data on the total photographic absorption A_p and the selective absorption C_e in some direction. Extra-galactic nebulae can be used for this purpose. We have seen how the counts of these nebulae give the optical half-thickness of the whole galactic absorbing layer in the direction perpendicular to the galactic plane. However, direct observations of the colours of extra-galactic objects (or of galactic objects lying outside the absorbing layer) can

give us the "selective" optical half-thickness of the absorbing layer, i. e. the difference $\tau_{\lambda_1} - \tau_{\lambda_2}$. Such a determination of the selective half-thickness has been made, for example, by the Soviet astronomer N. F. FLORYA, using the colour indices of globular clusters found by STEBBINS and WHITFORD. It is evident that we can obtain c_2/c_1 from the ratio of the selective half-thickness to the half-thickness in one of the wavelengths used:

$$\frac{\tau_{\lambda_1} - \tau_{\lambda_2}}{\tau_{\lambda_1}} = \frac{c_1(\lambda_1^{-1} - \lambda_2^{-1})}{c_1 \lambda_1^{-1} + c_2} = \frac{\lambda_1^{-1} - \lambda_2^{-1}}{\lambda_1^{-1} + c_2/c_1}.$$

It has been found that the observations are in accordance with the assumption that c_2 can be neglected, so that, within the limits of accuracy involved, they can be satisfied by the formula

$$\alpha(\lambda) = c_1 \lambda^{-1}. \tag{35.33}$$

Conversely, if we assume (35.33), it is possible to find, for any system of colour indices based on two given effective wavelengths, the factor γ whereby the directly observed colour excesses can be transformed into the values of the total absorption, in the photographic region, as far as the object concerned:

$$\gamma = \frac{c_1 \lambda_p^{-1}}{c_1 \lambda_1^{-1} - c_1 \lambda_2^{-1}} = \frac{\lambda_p^{-1}}{\lambda_1^{-1} - \lambda_2^{-1}}.$$

This has made it possible to use the comprehensive existing data on the photoelectric and photographic colour indices in order to ascertain the nature of the distribution of the absorbing matter in the Galaxy.

From a study of these data, P. P. PARENAGO has shown that previous estimates of the mean absorption coefficient in the galactic plane were too low, since the authors who made them used stars with galactic latitudes which, though small, were not zero, and they did not take into account the decrease in the absorption coefficient as we move away from the galactic plane. Having established that the "equivalent" linear half-thickness of the galactic absorbing layer is $\beta = 100$ parsecs, and taking as an interpolation formula an exponential law of variation of the density of the layer with altitude, we find that, if a_0 is the (photographic) absorption coefficient in the plane of the equator, and r and b are the distance of the object and its galactic latitude, then the total photographic absorption of the light of this object is

$$A(r,b) = (a_0 \beta/\sin b)\,(1 - e^{-r \sin b/\beta}), \tag{35.34}$$

and only in the particular case of an object for which $b = 0$ exactly do we have

$$A = a_0 r. \tag{35.35}$$

Hence the authors who have applied formula (35.35) directly to objects in latitude b have found, not the coefficient a_0, but the quantity

$$(a_0\,\beta/r \sin b)\,(1 - e^{-r \sin b/\beta})\,,$$

which is close to a_0 only if $r \sin b \ll \beta$.

The correct statement of the problem has led to mean values of a_0 for various directions as follows:

$a_0 = 3^{\mathrm{m}} \cdot 4$ per kiloparsec (PARENAGO),
$\quad = 2^{\mathrm{m}} \cdot 4$ per kiloparsec (FLORYA).

We note that N. F. FLORYA takes $\beta = 130$ parsecs.

Thus a considerable divergence is found from the value of a_0 obtained from formula (35.9). Since the latter formula assumed that j/α is everywhere the same as in the neighbourhood of the Sun, the discrepancy is to be explained, at least in part, by the fact that j/α at large distances from the Sun (thousands of parsecs or more) is, in some regions of space, large compared with its value near the Sun (the latter is in a region of the Galaxy lying between the spiral arms, where stars are rare).

In conclusion, we must particularly emphasise the result which has been reached by all investigators of the absorption in interstellar space, namely that the absorbing matter is distributed extremely irregularly, and the absorption coefficients obtained for different directions are themselves different.

2. **Diffuse nebulae.** We have already remarked above that the value of the absorption coefficient, even in the galactic plane itself, varies considerably from one direction to another, amounting to four stellar magnitudes per kiloparsec in some directions. The impression obtained from this of the extreme irregularity of the spatial distribution of the absorbing matter is reinforced by the facts relating to dark nebulae.

It is well known that there are certain regions where dark areas are seen against the bright background of the Milky Way. Such, for example, are the dark nebula around ϱ Ophiuchi, the Coal Sack in the southern sky, and many other objects. Similar regions, which are sometimes very small in diameter (less than a degree), are found, for example, in Cassiopeia and in Cygnus. Star counts on photographs show an extreme paucity of stars in these regions, and especially of faint stars. Since it is unlikely that this paucity of stars is due to an anomalously low density of stars within the cone formed by the lines joining the observer to the boundaries of this region, because this would mean assuming the existence of empty tunnels, almost devoid of stars and sometimes very narrow, in the stellar system, we

must draw the only possible conclusion, that there is absorbing matter near the observer in the direction concerned, which strongly attenuates the light of distant stars. The absorption, i. e. the optical thickness, of such **dark nebulae**, and their distance from us, can be approximately estimated. In some cases the optical thickness amounts to two or three stellar magnitudes. In order that a large dark nebula should be fairly distinctly seen against the background of the Milky Way, its distance must be *not more than four hundred parsecs.* At greater distances the number of stars which are nearer, and whose light is unaffected by the nebula, is so large that we cannot speak of an exceptional paucity of stars observed in the region of the nebula.

The cross-sectional diameters of dark nebulae can be estimated from their distances and angular dimensions; they are found to be of the order of ten parsecs. There are, in fact, more dark nebulae with smaller cross-sectional dimensions, but they are less noticeable. It is clear also that the linear extent in depth is usually of the same order as the cross-sectional dimensions, since there is no reason to suppose that all the dark nebulae are elongated in the direction of the line of sight.

The dark nebulae are similar to the "globules", which are dark regions of the sky with angular diameters of the order of a minute or less, and are mostly found against the background of bright diffuse nebulae.

Apart from the globules, whose nature is as yet by no means clear, it must be supposed that the dark nebulae, like all the absorbing matter in the Galaxy, consist of solid particles of cosmic dust. Since observation reveals some selectivity in their absorbing properties, we must conclude that the diameters of these particles do not exceed the wavelength in the majority of cases. Using more or less plausible hypotheses concerning the density of these particles, we find a mass of the order of $50\ M_\odot$ for the largest of the dark nebulae. The dark nebulae of smaller dimensions have masses not exceeding $10\ M_\odot$.

The shape of the regions screened by the dark nebulae is usually quite irregular. Moreover, the boundaries of these regions are fairly indefinite. These facts indicate that the dark nebulae themselves usually have irregular diffuse shapes. This links them to the **bright diffuse nebulae** having a continuous spectrum, the investigation of which has shown that they also are clouds of cosmic dust, like the dark nebulae, the only difference being that they are illuminated by stars of high luminosity.

It is well known that the bright diffuse nebulae which we observe are divided into two groups. Some of them have spectra composed of bright lines of various light elements (**nebulae with emission spectra**) whilst the others (**reflecting nebulae**) exhibit *continuous spectra with ab-*

sorption lines, identical with those of stars of one of the familiar spectral types. The question has naturally arisen of the source of the radiation of each type of diffuse nebula.

In the spectra of diffuse nebulae with emission, the same lines are found as in the spectra of the planetary nebulae. The chief difference lies in the fact that, in the latter spectra, the N_1 and N_2 lines of O III are dominant, being much more intense than the adjacent H_β line. In the spectra of the diffuse nebulae with emission, however, they are not brighter than H_β. The ratio of brightnesses of the lines of O III and O II is also much less than in the case of the planetary nebulae. This, together with the absence of lines of ionised helium in the spectra of the diffuse nebulae, shows that the difference between the two cases lies only in the degree of excitation and ionisation of the atoms, which is less for the diffuse nebulae. On the other hand, the features of similarity in the spectra indicate that the mechanism of excitation of the spectra is the same in each case, i. e. in the case of the diffuse nebulae also the spectra are excited by the ultra-violet radiation of a hot star. E. P. HUBBLE has shown that in fact, close to almost every diffuse nebula with emission, a giant star of class O or B 0 can be found, or else a group of such giants, which can be regarded as exciting the radiation in accordance with the mechanism of which the theory has been explained in detail in Part IV, in connection with the planetary nebulae. An attempt to determine the temperatures of these giants by ZANSTRA's method, on the assumption that they are in fact the source of the luminosity of the corresponding diffuse nebulae, has led to estimates which agree well with present-day views on the temperatures of stars of classes O and B 0 (30,000° to 25,000°), as obtained by other methods. Thus each diffuse nebula with emission is illuminated by a hot star or by a group of hot stars which cause it to radiate. There are only a few exceptions to this statement.

In the case of diffuse nebulae having a continuous spectrum, there may always be found, near the nebula, a supergiant of spectral type later than B 0, whose spectrum is similar to that of the nebula (if the latter is known). In this case it can be assumed that the nebula, which consists of solid particles, reflects the light from this supergiant, or, more precisely, scatters it.

HUBBLE has shown that, if we determine, for each diffuse nebula, the greatest distance a from the illuminating star (i. e. from the supergiant mentioned above) to points of the nebula having some previously fixed standard surface brightness, it is found empirically that a is related to the apparent magnitude m of the illuminating star by

$$m + 5 \log_{10} a = \text{constant}. \tag{35.36}$$

If we construct a graph of m against $\log_{10} a$, the points corresponding to the individual nebulae are concentrated around the straight line (35.36) with only a small dispersion. On the other hand, it can be shown that the relation (35.36) follows immediately from the hypothesis that the diffuse nebulae are illuminated by stars. Thus, let us denote by b the light from the star which reaches the Earth. Here we refer to the illumination of areas perpendicular to the star's rays. The illumination of a dust nebula whose particles lie in some plane perpendicular to the direction of the star's rays is

$$b\, r^2/\Delta^2 ,$$

where r and Δ are respectively the distances from the star to the Earth and to the point in the nebula considered. If the reflecting power of the nebula is constant (for example, if it reflects practically all the light from the star), the observed surface brightness is proportional to the illumination, i. e. $b\, r^2/\Delta^2$. Moreover, as is well known, it is independent of the distance between the nebula and the observer (neglecting absorption).

Consequently, if we select points, in various nebulae, which correspond to some standard surface brightness, the distance Δ at these points satisfies the equation

$$\Delta^2 = Cbr^2 , \tag{35.37}$$

where C is a constant. On the other hand, we have for the angular distance a from the star to the point considered in the nebula

$$a = (\Delta/r) \cos \theta , \tag{35.38}$$

where θ is the angle made by the radius vector drawn from the star to the point in question with the tangent plane to the sky. From (35.37) and (35.38) we obtain

$$a^2 = C\, b \cos^2 \theta ,$$

or, taking logarithms,

$$2 \log_{10} a - \log_{10} b = \log_{10} C + 2 \log_{10} \cos \theta .$$

Multiplying by 2·5 and taking into account the relation

$$- 2{\cdot}5 \log_{10} b = m + C' ,$$

where m is the apparent magnitude of the star, we find

$$5 \log_{10} a + m - 5 \log_{10} \cos \theta = \text{constant} . \tag{35.39}$$

The angle θ is usually unknown, and the relation (35.39) cannot be exactly tested. However, replacing $\log_{10} \cos \theta$ by its mean value, we obtain equation (35.36), which is in good agreement with observation. The slight dispersion of the points, on the graph of m against $\log_{10} a$, about the straight line (35.36) is due, firstly, to the fact that $\log_{10} \cos \theta$ differs from its mean value in different nebulae and, secondly, to the fact that different nebulae have different albedos.

The approximate fulfilment of equation (35.36) not only confirms the hypothesis that the radiation of the diffuse nebulae with a continuous spectrum is the result of the scattering by the nebulae of the light of the illuminating stars, but also shows that the illuminating supergiant stars have, in the majority of cases, been correctly identified by astronomers. This is reasonable, since the number of supergiants in the Galaxy is relatively small, and the identity of the spectral types of star and nebula is a criterion which makes improbable any error in the identification of the illuminating stars.

The proved fact that in this case we are concerned with the scattering of the light of the star by a dust nebula indicates that the physical state of the matter in the dark nebulae is the same as in the bright diffuse nebulae with a continuous spectrum.

It is interesting that the relation (35.36) is fulfilled, not only for diffuse nebulae with a continuous spectrum, but also for emission nebulae. This fact is somewhat unexpected, since in this case we do not have simple reflection, but, as was said above, a process similar to that which occurs in the planetary nebulae. The fulfilment of the relation (35.36) for the emission nebulae indicates that, from unit surface of the nebula, an amount of energy is emitted in all the emission lines together which is proportional to the amount of energy incident on this unit surface in the usual photographic part of the spectrum. We know, however, that the amount of energy emitted in the spectral lines by unit surface of a nebula depends on the amount of radiation from the star, beyond the limit of the Lyman series, that is incident on 1 cm^2 of the surface of the nebula.

Hence the fulfilment of the relation (35.36) indicates that the ratio of the amount of energy incident on 1 cm^2 of the nebula in photographic wavelengths to the amount of energy incident on the same area in the far ultra-violet region of the spectrum is approximately constant from one nebula to another. This, however, means that the surface temperatures of the illuminating stars are the same. We know that the stars which illuminate the diffuse nebulae having an emission spectrum do in fact belong to a very narrow range of the spectral sequence, namely the classes O and B 0 with absorption lines. As a rule, no diffuse nebulae are found which are illuminated by Wolf-Rayet

type stars. Thus the temperatures of the stars which illuminate the diffuse nebulae having an emission spectrum lie in the interval from 20,000° to 30,000°. It is by chance that the constant on the right-hand side of equation (35.36) is approximately the same for reflecting and emission nebulae. These constants would be different if the temperatures of the stars which cause the radiation of the emission nebulae were different.

The distances of the nebulae from the stars which illuminate them is often several parsecs. The question arises whether there is a *genetic relation* between the diffuse nebulae and the illuminating stars, or whether their association is purely casual. In other words, do the star and the nebula have a common origin and motion, or have they approached by chance and come close together at the present time, only to separate again? HUBBLE himself, on comparing the radial velocities of several nebulae and the stars illuminating them, noticed discrepancies between them, and hence concluded that their relation is purely temporary. HUBBLE's conclusion, however, was based on the radial velocities of only five stars, and needed to be carefully tested. This has been carried out by SH. G. GORDELADZE and V. A. AMBARTSUMYAN in the following manner.

If the nebulae meet their illuminating stars only by chance, the number of nebulae illuminated at any time by the stars of some spectral class should be proportional to the probability that a nebula is within the volume illuminated by one or other of the stars of the spectral type in question. This is due to the obvious circumstance that each star having a given luminosity can produce an illumination exceeding some limiting value only if the object illuminated is within some spherical volume. If, in addition, the reflecting power is given, we can draw round each star a sphere such that, if the nebula is within the sphere, it will have a surface brightness exceeding some definite value. Here we can take as the minimum surface brightness the value detected by astronomical instruments of a given focal ratio (say 1 : 5) for a given exposure (say one hour) and a given plate sensitivity. The radius of this sphere is proportional to the square root of the star's luminosity, i.e.

$$\log_{10} \Delta = C - 0{\cdot}2\, M\,,$$

where M is the absolute magnitude of the star. The volume of the sphere is proportional to the cube of the radius, i.e.

$$\log_{10} V = C_1 - 0{\cdot}6\, M\,. \tag{35.40}$$

Thus, for example, calculation shows that, for the above conditions of observation, stars of absolute magnitude $M = 0$ illuminate about 1 cubic parsec around them, while those of absolute magnitude $M = -5$

illuminate about 1000 cubic parsecs. Outside these volumes, the nebula would not be accessible to instruments of the given focal ratio with an hour's exposure, since its surface brightness would be too small.

It is evident also that the fraction of the total volume illuminated by all the stars of some spectral type is $n\,\bar{V}$, where $\bar{V}$ is the mean volume illuminated by a star of that type and n is the number of stars of that type per unit volume. Here the mean volume $\bar{V}$ can be obtained by averaging, over all values of the absolute magnitude M, the volume $V(M)$ calculated from formula (35.40):

$$\bar{V} = \int_{-\infty}^{\infty} \phi(M)\; V(M)\, \mathrm{d}M \Big/ \int_{-\infty}^{\infty} \phi(M)\, \mathrm{d}M\,, \tag{35.41}$$

where $\phi(M)$ is the luminosity function for a given spectral type, which shows the fraction of stars of the type in question which have absolute magnitudes between $M - \frac{1}{2}$ and $M + \frac{1}{2}$. Thus the probability that a nebula is illuminated by a star of the given spectral type can be calculated from data of stellar statistics only. If the relation between the nebulae and the stars illuminating them is in fact one of chance, the number of observed nebulae illuminated by stars of some type should be proportional to $n\,\bar{V}$. Table 35 gives, for each spectral type, the number N of observed nebulae illuminated by stars of that type, according to HUBBLE, and the fraction $n\,\bar{V}$ of galactic space illuminated by stars of the same type, as calculated from formula (35.41) and from the data of stellar statistics concerning the luminosity functions and the frequency of occurrence of the various types of stars. We see that the hypothesis of a chance relation excellently explains the predominance, among the reflecting nebulae, of those associated with stars of types B 1 to B 9, and is generally in good agreement with observation.

Table 35

Spectral class	N	$n\,\bar{V}$	p
O	11	$0{\cdot}2\times10^{-4}$	$0{\cdot}2\times10^{-4}$
B 0	7	0·6	0·2
B 1 – B 9	54	2·9	2·2
A	5	0·8	0·6
F	2	0·25	0·5
G	1	0·18	0·1
K	2	0·25	0·4
M	0	0·02	0·05

We note that, instead of using the luminosity functions for the various spectral types in accordance with (35.41), the results for which are still insufficiently reliable, the same quantities $n\,\bar{V}$ can be obtained

as follows. In our calculations, we have not taken into account the change in the numbers n as we move away from the Sun along the galactic plane. Hence we can bring the constant n under the integral sign in the formula

$$p = n \int_{-\infty}^{\infty} V_0 \cdot 10^{-0\cdot 6 M} \phi(M) \, \mathrm{d}M ,$$

which, by (35.40) and (35.41), gives the fraction of a unit volume which is illuminated by stars of a given type. Here V_0 is the volume illuminated by a star of absolute magnitude zero. Thus

$$p = V_0 \int_{-\infty}^{\infty} n \, \phi(M) \cdot 10^{-0\cdot 6 M} \, \mathrm{d}M . \tag{35.42}$$

On the other hand, according to the fundamental integral equation of stellar statistics, we have for the number of stars of apparent magnitude m in any direction in the galactic plane

$$A(m) = \omega \int_{0}^{\infty} n \, \phi(M) \, r^2 \, \mathrm{d}r . \tag{35.43}$$

Expressing r by means of

$$5 \log_{10} r - 5 = m - M ,$$

we can rewrite (35.43) in the form

$$A(m) = \frac{\omega}{5 M} \int_{-\infty}^{\infty} n \cdot 10^{0\cdot 6(m-M)+3} \phi(M) \, \mathrm{d}M . \tag{35.44}$$

The absorption of light in interstellar space has not been taken into account in formulae (35.43) and (35.44). However, if we restrict ourselves to stars above the fifth magnitude, i.e. to the nearest stars, this is quite permissible. Then, comparing (35.44) and (35.42), we find that

$$A(m) = (\omega/5 M) \, 10^{0\cdot 6 m+3} \, p ,$$

from which the values of p can be calculated, using the values of $A(m)$ for various spectral types with any m. Thus we are enabled to obtain these values directly from observation. The values of p thus found are given in the last column of Table 35.

As we see, the relative values of p in this case also correspond to the relative values of the observed numbers of diffuse nebulae. Thus *the statistics of the spectra of the illuminating stars is in good agreement with*

the hypothesis of casual connection. We notice that, for the types O and B 0, where we are concerned with gaseous nebulae, despite the presence of some deviations the same agreement with observation is obtained. However, as we shall see below, the problem of the gaseous nebulae is more complex than appears at first sight. Hence we shall at present draw some conclusions regarding only the dust nebulae.

The most important result which follows from the confirmation of the hypothesis of casual connection is that, besides the dust nebulae illuminated by stars, there must also exist in the stellar system a very great number of nebulae which are not illuminated.

In fact, the numbers p denote the fractions of any volume of galactic space which are illuminated by stars of a given spectral class. The sum of all the p, however, not only is not equal to unity, but is only of the order of 1/2000. In other words, the stars of all the spectral classes together illuminate only one two-thousandth part of interstellar space. It then follows, from our conclusion regarding the casual connection between the dust nebulae and the stars illuminating them, that the total number of nebulae must be approximately two thousand times the number of illuminated nebulae. Consequently, only a few representatives of the dust nebulae are illuminated by stars.

On the other hand, just as, among the illuminated diffuse nebulae, there are objects of various sizes and brightnesses, among the nebulae that are not illuminated there are objects of various diameters and optical thicknesses.

Those which are not far from us and have particularly large optical thicknesses will attenuate the light of the stars lying behind them. and produce regions of the sky in which there are few stars. We observe them as "dark" nebulae. Thus we see that both the bright and the dark nebulae are particular cases of a very numerous class of diffuse nebulae.

If we take into account the fact that, with instruments of focal ratio 1 : 5 and exposures of one hour, upwards of a hundred illuminated diffuse nebulae can be photographed with present-day plates, and that the great majority of these are at distances from us of less than 1500 parsecs and lie within a galactic layer of thickness 200 parsecs, we must suppose that upwards of two hundred thousand unilluminated nebulae actually lie in this volume (since $\Sigma\, p = 1/2000$). Assuming that the radius of each nebula is of the order of three parsecs, we can calculate the number of such nebulae which intercept a ray in the galactic plane over a distance of 1000 parsecs. We obtain a number of the order of four nebulae per kiloparsec of light path. Since these nebulae (or clouds, as we shall now call them) must have a certain absorbing (or scattering) power, as is shown, in particular, by the considerable albedo which they exhibit when illuminated, we must suppose that each of these clouds brings about

a noticeable attenuation of the light of the stars lying behind it, i.e. has a considerable optical thickness. If we further assume that this thickness, expressed in stellar magnitudes, is about $0^{\mathrm{m}}\cdot 3$, it is found that the non-illuminated clouds together should bring about an absorption of more than 1^{m} per kiloparsec. Thus we can draw the conclusion that the general and selective absorption of which we spoke in the last section is, if not entirely, at least in part the total effect of the absorbing clouds whose existence has been established above. We shall discuss below the consequences of assuming that the absorbing layer consists entirely of an assembly of separate clouds. The case where part of the absorption is due to a continuous medium can be reduced to this, since we can suppose that the continuous medium has the same effect as an infinite (or very large) number of clouds, each of infinitely small (or very small) optical thickness.

The discrete structure of the absorbing layer causes certain deviations from uniformity in the apparent distribution of stars and of extra-galactic nebulae. We shall begin by considering the irregularities in the distribution of extra-galactic nebulae.

3. Fluctuations in the numbers of extra-galactic nebulae. We have seen above that, if we assume the absorbing matter to be distributed in the Galaxy in continuous plane-parallel layers parallel to the plane of the galactic equator, the number of extra-galactic nebulae down to some apparent magnitude m in one square degree is a function only of galactic latitude. Thus in these conditions N_m should not vary for a given galactic latitude. However, if the absorbing layer consists of separate clouds, the number of clouds intersected by rays from two neighbouring directions, even with the same galactic latitude, will in general be different. Hence the absorption in these two directions will be different, and consequently N_m will be different also.

As before, we shall start from the supposition that, if there were no absorption in the Galaxy, N_m would be the same in all directions, namely

$$N_m = N_0 \cdot 10^{0{\cdot}6\, m_0}\,, \tag{35.45}$$

where m_0 is the apparent magnitude that would be observed in the absence of absorption. Let us consider rays reaching us from extra-galactic nebulae lying in some direction at galactic latitude b. After entering the Galaxy, they traverse a series of absorbing clouds, as a result of which their intensity is diminished by a factor Q. It is evident that

$$Q = e^{-\tau}\,, \tag{35.46}$$

where τ is the total optical thickness of all the clouds intersected by the ray. Then the nebulae which would have an apparent magnitude m_0 if there

were no absorption in the Galaxy will be observed with an apparent magnitude

$$m = m_0 - 2{\cdot}5 \log_{10} Q\,. \tag{35.47}$$

Substituting the expression for m_0 from (35.47) in (35.45), we find

$$N_m = N_0 \cdot 10^{0\cdot 6\, m}\, Q^{3/2}\,. \tag{35.48}$$

The quantity Q is a random variable. We can study the statistical law of distribution of various values of N_m at a given galactic latitude, using the observed values of N_m in various directions. Hence it is convenient to form an idea of the properties which this distribution should have as a function of the statistical parameters characterising the absorbing clouds.

Let us rewrite (35.48) by introducing the value of Q from (35.46). We then have

$$N_m = N_0 \cdot 10^{0\cdot 6\, m}\, e^{-3\tau/2}\,. \tag{35.49}$$

Thus the law of distribution of N_m depends on that of τ, i.e. on the total optical thickness traversed by the ray. The optical thickness τ traversed by the ray is composed of the optical thicknesses traversed by it in the separate clouds:

$$\tau = \Sigma\, \sigma_i\,, \tag{35.50}$$

where both the number of terms in the sum and the optical thickness traversed by the ray in each cloud, i.e. the value of each σ_i, are random quantities. This is due to the fact that the clouds may differ among themselves; moreover, the ray may intersect a cloud along lines corresponding to various values of the optical thickness.

Let $F(\sigma)$ be the probability that a cloud intersected by the ray has an optical thickness less than σ along the line of intersection. We suppose that the function $F(\sigma)$ is independent of the distance from the galactic plane.

We have to find the law of distribution of τ (the optical thickness of the whole layer of clouds in a given latitude, from the observer to the outer boundary of the layer). However, in order to solve this problem, we shall take at first only the part of the cloud layer up to an altitude z from the galactic plane, and then increase z. We denote by t the optical thickness of this layer in the direction of latitude b, and desire to find the distribution function of t, i.e. the probability $\phi(t)$ that the optical thickness traversed by the ray in this layer is less than t.

Let us add to the layer of clouds some linear increment of thickness $\mathrm{d}z$. Then the path traversed by a ray in this layer is increased by the segment $\mathrm{d}z \operatorname{cosec} b$. Here two cases are possible: either a cloud is encountered in this segment, or it is not. The probability of finding a cloud is proportional to the length $\mathrm{d}z \operatorname{cosec} b$ of the segment, and is $k\,\mathrm{d}z \operatorname{cosec} b$, where k is the mean number of clouds intersected in unit path of the ray at the given altitude z. The probability of not finding a cloud is $1 - k\,\mathrm{d}z \operatorname{cosec} b$. The distribution functions $\phi(t)$ before and after the increment will differ by their parameters z and $z + \mathrm{d}z$. Then, by the theorems of composition and multiplication of probabilities, we have

$$\phi_{z+dz}(t) = (1 - k\,\mathrm{d}z \operatorname{cosec} b)\,\phi_z(t) + k\,\mathrm{d}z \operatorname{cosec} b \int \phi_z(t-\sigma)\,\mathrm{d}F(\sigma)\,. \quad (35.51)$$

Transferring $\phi_z(t)$ to the left-hand side and dividing by $k\,\mathrm{d}z \operatorname{cosec} b$, we obtain the differential equation

$$\sin b\,\frac{\mathrm{d}\phi}{k\,\mathrm{d}z} = -\phi_z(t) + \int_0^t \phi_z(t-\sigma)\,\mathrm{d}F(\sigma)\,. \quad (35.52)$$

We introduce instead of z the independent variable

$$v = \int_0^z k\,\mathrm{d}z\,. \quad (35.53)$$

Then equation (35.52) can be rewritten in the form

$$\sin b\,\frac{\mathrm{d}\phi_v}{\mathrm{d}v} = -\phi_v(t) + \int_0^t \phi_v(t-\sigma)\,\mathrm{d}F(\sigma)\,. \quad (35.54)$$

From this equation it is easy to obtain the expectations of various powers of t. However, in order to proceed with the problem of the fluctuations of the numbers of extra-galactic nebulae, we do not need these expectations, but those of various powers of $e^{-3\tau/2}$. To calculate these, we multiply equation (35.54) by e^{-lt} and integrate from 0 to ∞. Then

$$\sin b\,\frac{\mathrm{d}}{\mathrm{d}v}\int_0^\infty e^{-lt}\,\phi_v(t)\,\mathrm{d}t =$$

$$-\int_0^\infty e^{-lt}\,\phi_v(t)\,\mathrm{d}t + \int_0^\infty e^{-lt}\,\mathrm{d}t \int_0^t \phi_v(t-\sigma)\,\mathrm{d}F(\sigma)\,.$$

By changing the order of integration we obtain, after a slight transformation of the second term on the right-hand side,

$$\sin b \frac{\mathrm{d}}{\mathrm{d}v} \int_0^\infty e^{-lt} \phi_v(t) \,\mathrm{d}t =$$

$$-\int_0^\infty e^{-lt} \phi_v(t) \,\mathrm{d}t + \int_0^\infty e^{-l\sigma} \,\mathrm{d}F(\sigma) \int_0^\infty \phi_v(x) \, e^{-lx} \,\mathrm{d}x \,. \qquad (35.55)$$

But

$$\int_0^\infty e^{-lt} \phi_v(t) \,\mathrm{d}t = -\frac{1}{l} e^{-lt} \phi_v(t) \Big|_0^\infty + \frac{1}{l} \int_0^\infty e^{-lt} \,\mathrm{d}\phi_v(t) = \frac{1}{l} [e^{-lt}] \,,$$

since $\phi(0) = 0$ and $\phi(\infty) = 1$. Hence we obtain from (35.55), on multiplying by l,

$$\sin b \,(\mathrm{d}[e^{-lt}]/\mathrm{d}v) = -[e^{-lt}] + [e^{-l\sigma}]\,[e^{-lt}] \,. \qquad (35.56)$$

Here, and in future, we denote the expectations of quantities by placing square brackets round them. The quantity

$$q = e^{-\sigma} \qquad (35.57)$$

is the *transparency* of one cloud. Hence

$$[e^{-l\sigma}] = [q^l]$$

denotes the mean value of the lth power of the transparency of one cloud when it is randomly oriented with respect to the ray traversing it. We also put

$$[e^{-lt}] = g_v(l) \,. \qquad (35.58)$$

Then equation (35.56) can be rewritten in the form

$$\sin b \,\mathrm{d}g_v(l)/\mathrm{d}v = -g_v + [q^l]\, g_v \,. \qquad (35.59)$$

Integration of this equation gives

$$g_v(l) - C\, e^{-v(1-[q^l])/\sin b} \,. \qquad (35.60)$$

Since, for $v = 0$, i.e. $z = 0$, $g_v(l) = 1$, we have $C = 1$. Hence

$$g_v(l) = e^{-v(1-[q^l])/\sin b} \,. \qquad (35.61)$$

In particular, we can apply this formula to the limiting value of v when z increases and we have the entire layer of clouds. We denote this value of v by v_0. Then we have for the expectation of $e^{-l\tau}$

$$g_{v_0}(l) = e^{-v_0(1-[q^l])/\sin b} . \tag{35.62}$$

This result can now be used to calculate the expectations of various powers of N_m. According to (35.49) we have

$$[N_m{}^k] = N_0{}^k \cdot 10^{0\cdot 6\,km}\,[e^{-3k\tau/2}] = N_0{}^k \cdot 10^{0\cdot 6\,km}\,g_{v_0}(3\,k/2) . \tag{35.63}$$

Hence, from (35.62),

$$[N_m{}^k] = N_0{}^k \cdot 10^{0\cdot 6\,km}\,e^{-v_0(1-[q^{3k/2}])/\sin b} . \tag{35.64}$$

In particular, we have for the mean value of N_m

$$[N_m] = N_0 \cdot 10^{0\cdot 6\,m}\,e^{-v_0(1-[q^{3/2}])/\sin b} , \tag{35.65}$$

and for the mean square deviation

$$[(N_m - [N_m])^2] = N_0{}^2 \cdot 10^{1\cdot 2\,m}\,\{e^{-v_0(1-[q^3])/\sin b} - e^{-2v_0(1-[q^{3/2}])/\sin b}\} . \tag{35.66}$$

In calculations of this type the *relative* value of the mean square deviation, i.e.

$$[(N_m - [N_m])^2]/[N_m]^2 = e^{v_0[(1-q^{3/2})^2]/\sin b} - 1 \tag{35.67}$$

is of greatest interest.

Formula (35.65) shows that the relation between the observed values of $\log_{10}[N_m]$ and $\operatorname{cosec} b$ should be linear. After finding this relation from observational data, we can determine the slope of the line representing it, i.e. the numerical value of

$$f_1 = v_0\,[1 - q^{3/2}] . \tag{35.68}$$

Having found from observation the mean square deviation from the mean over various galactic latitudes, we can also determine, from (35.67), the numerical value of

$$f_2 = v_0\,[(1-q^{3/2})^2] . \tag{35.69}$$

However, in the two equations (35.68) and (35.69) three quantities appear which characterise the assembly of absorbing clouds and are to be determined. These are v_0, $[1 - q^{3/2}]$ and $[(1 - q^{3/2})^2]$. It is evident that we cannot determine them all at once. If we assume, however, that

the values of q, for instance, have only a small dispersion, i.e. that all the values of the transparency q are equal to the same value q_0, we find that the third of these quantities is equal to the square of the second, and so all three are determined. Let us find q_0 and v. In other words, let us determine q_0 from the relation

$$f_2/f_1 = [(1 - q^{3/2})^2]/[1 - q^{3/2}] = 1 - q_0^{3/2} . \qquad (35.70)$$

Now the mean value of a square is always greater than the square of the mean value. Hence

$$1 - q_0^{3/2} = [(1 - q^{3/2})^2]/[1 - q^{3/2}] > [1 - q^{3/2}]$$

or

$$[q^{3/2}] > q_0^{3/2} .$$

Thus the $q_0^{3/2}$ determined above, which can be calculated directly from observational data, is less than the mean value of $q^{3/2}$.

Although, as we see, the q_0 calculated according to (35.70) is not exactly equal to $[q]$ or to $([q^{3/2}])^{2/3}$, it nevertheless gives some idea of the value about which the actual values of q are concentrated. If we use, to find $[N]$ and $[N^2]$, the counts made by E. P. Hubble and H. Shapley, a value of the order of 0·8 is obtained for q, i.e. each cloud transmits 80 % of the light incident on it. Thus the attenuation of the starlight caused by an average individual cloud is usually small, and therefore, instead of noticing the effect of each cloud separately, we observe the total effect due to all the clouds, in the form of an absorption which increases with distance. Only if q deviated considerably from q_0 towards smaller values, reaching values of 0·2 or 0·3, should we observe the absorbing clouds in the form of individual dark nebulae.

We note that, if part of the absorbing matter is distributed in the form of a continuous medium, or consists of a large number of clouds of very small optical thickness, the probability is high that q, and consequently $q^{3/2}$, will be close to unity. To explain the observed value of f_2/f_1, it is then necessary to assume that the mean absorption in the remaining clouds is considerably greater than when the presence of such a continuous medium is not assumed in explaining this value of f_2/f_1.

The theory developed in this chapter takes no account of the natural fluctuations in the numbers of galaxies. If this is allowed for, there will be some decrease in the effective value of $[q]$. The "clustering" tendency of the galaxies may have a still greater effect. Since 1952 there has appeared a series of papers by C. D. Shane, J. Neyman and E. L. Scott, who consider a model of the spatial distribution of galaxies based on the hypothesis that all galaxies belong to one of a number of clusters

[113]. Recent results concerning the distribution of galaxies indicate that the clustering tendency has a very great effect. In this case, a considerable part of the apparent fluctuations must arise from this tendency, and the value of $[q]$ must be still smaller.

On the other hand, T. A. AGEKYAN [1] has shown that, starting from the hypothesis that the observed fluctuations are entirely due to variations in the galactic absorption, and treating the existing data more carefully, we find that the value of $[q]$ must be increased. It would be desirable to re-examine the existing observational data on the basis of a theory which takes account of both factors at the same time.

4. **The polarisation of starlight.** One of the most interesting astronomical discoveries in recent years has been that of the polarisation of the light from remote stars, made by W. A. HILTNER, by J. S. HALL, and by V. A. DOMBROVSKIĬ. The observations show that the stars in any particular small area of the sky usually have nearly the same plane of polarisation. A study of this phenomenon has led to the conclusion that it is due to the fact that the cause of the polarisation is interstellar. Since the degree of polarisation, at least at not very great distances from the observer, shows a clear correlation with the colour excess, the idea has been put forward that the polarisation is due to the same interstellar matter as causes the reddening of starlight.

In order that dust particles should be able to bring about such an anisotropic effect as polarisation, it is necessary that these particles should be extended or compressed, and also that there should be some agency causing a predominant orientation of the particles in some direction. This agency may be the action of a magnetic field, and this should cause also a magnetisation of the particles.

Two different theories have so far been advanced to account for the orientation of particles in a magnetic field. Both, however, meet with certain difficulties.

Although the existence of interstellar magnetic fields of intensity between 10^{-6} and 10^{-4} gauss is now undoubted, V. A. DOMBROVSKIĬ has discovered a number of interesting facts which indicate that, besides the interstellar mechanism, a local (perhaps circumstellar) mechanism also sometimes acts. For example, cases are known where the components of a binary star exhibit different polarisation. Similar results have been obtained by HALL.

In the region of the Milky Way, the electric intensity vector is very nearly in the plane of the Galaxy, and this is regarded as evidence that the magnetic lines of force are approximately parallel to this plane.

5. The general radio emission. The hypothesis that interstellar magnetic fields exist in the Galaxy has been very fruitful also in explaining the general radio emission of the Galaxy. V. L. GINZBURG has explained this radio emission at metre wavelengths by the well-known emission of electromagnetic waves by relativistic charged particles moving in a magnetic field.

Theoretical calculations show that the energy lost by a relativistic electron when it moves in a magnetic field H is given by the formula

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{2}{3}c\left(\frac{e^2}{m\,c^2}\right)^2 H^2 \left(\frac{E}{m\,c^2}\right)^2,$$

where E is the energy of the electron, and the other quantities are known universal constants.

The intensity of the radiation emitted by particles of a given energy E increases with frequency for small frequencies, reaches a maximum, and then decreases exponentially. Thus the spectrum is limited quite sharply on the short-wave side. The frequency of the maximum is given by

$$\nu = \frac{1}{4\pi}\,\frac{eH}{m\,c}\left(\frac{E}{m\,c^2}\right)^2.$$

If we take the magnetic field to be of the order of 10^{-5} oersted, then the emission of the Galaxy in the metre wavelength range can be explained by supposing that there exist in the interstellar medium electrons with energies of the order of 10^9 eV.

The hypothesis that there are relativistic electrons in interstellar space might in itself seem artificial, were it not for the fact that cosmic rays furnish us with a direct proof of the presence there of a large number of high-energy particles.

On the above theory, the radio emission of the Galaxy is non-thermal and non-equilibrium. The high temperature of the radio emission of the Galaxy in the metre wavelength range has always been regarded as a demonstration of non-equilibrium in the mechanism that gives rise to it.

Chapter 36. The gaseous component of interstellar matter

1. The interstellar gas. Besides the general attenuation of the light of the stars, as a result of the absorption of their radiation by interstellar dust in all the frequencies of the visible and photographic regions of the spectrum, i. e. besides the "continuous" absorption in

interstellar space, we often observe in the spectra of distant stars absorption lines due to interstellar matter in the *gaseous* state. In the parts of the spectrum accessible to study, the lines which attract attention are, first of all, the interstellar lines of Ca II (the H and K lines), and the yellow doublet 5890 and 5896 Å and the ultra-violet doublet 3302 and 3303 Å of Na I. Moreover, interstellar lines of Ca I, K I, Ti II and the molecules CH I and CH II are observed. In every case we are concerned with lines absorbed by an atom, ion or molecule in the ground state. This is understandable, since, in consequence of the extreme dilution of the radiation in interstellar space, the number of atoms in an excited state must be extremely small.

Since the observation of the interstellar lines is hampered by the absorption lines formed in the star's atmosphere, the presence of the former can as yet be detected only for stars in whose spectra this interference is least. The presence of absorption lines due to matter outside the star's atmosphere was first discovered by studying the radial velocities of some hot giants which are components of spectroscopic binaries. It was found that the H and K lines in their spectra have components exhibiting a *constant* velocity, while all the other lines show a periodic variation of velocity in consequence of the orbital motion. These components of the Ca II lines differ from the stellar components of the same lines by their external form also: they are much narrower and sharper. The idea thus arose that these lines are produced outside the atmospheres of the two components of the binary.

Later, similar narrow absorption lines of Ca II were discovered in the spectra of single stars of classes O and B, and it was found that the radial velocities determined from these lines differ from those determined from the other stellar lines. This confirmed the hypothesis that there exists an interstellar gas in which there are Ca II ions. Similar conclusions were drawn from an investigation of the yellow sodium doublet D_1 and D_2 ($\lambda\lambda$ 5890 Å, 5896 Å).

The fact that these interstellar calcium lines are observed only for stars of classes O and B gives no reason to suppose that the gaseous interstellar matter which produces these lines is concentrated around these stars only. For stars of the later types have in their spectra intense and very broad H and K absorption lines of their own, and this makes it impossible to observe the faint interstellar calcium lines superposed on them. Hence it is important to take into consideration the fact that the interstellar sodium doublet $D_{1,2}$ can be observed, not only in the spectra of stars of classes O and B, but also in those of remote stars of class A (principally supergiants, which can be observed at great distances). This has enabled us to draw the conclusion that the interstellar gas is distributed everywhere in the galactic plane, forming a layer on

both sides of this plane. This conclusion has been entirely confirmed by observations of distant stars in low galactic latitudes, in whose spectra stronger interstellar absorption lines are found. In other words, a correlation has been found between the distances of stars of classes O and B and the equivalent widths of the interstellar lines. However, it has been found that, although the intensity of the interstellar lines increases on the average with distance, deviations from this are also observed. This fact has been correctly explained as being due to an irregularity in the distribution of the interstellar gas in different directions. The same irregularity is observed in the distribution of the interstellar gas as in that of the interstellar dust.

On the other hand, the irregularity in the distribution of the interstellar dust is, as we have seen above, expressed by saying that it is distributed in the Galaxy in the form of separate dust clouds. The question arises whether the layer of interstellar gas has the same patchy, cloudy structure: are not the deviations mentioned above largely accounted for by fluctuations in the number and dimensions of the gaseous clouds through which a ray of light passes from a distant star?

The investigation of the interstellar lines in the spectra of stars, by means of spectrographs of high dispersion, has provided the answers to these questions. It has been found that the interstellar lines in the spectra of distant stars often consist of several components, corresponding to the passage of the ray from the star through several clouds having different radial velocities. Thus the possibility has been revealed of knowing in each case the number of clouds through which the light of the star passes, although we must here bear in mind the possible coalescence of components due to different clouds when their radial velocities are similar. This is true, in particular, of directions where the differential effect of galactic rotation on the radial velocities is almost zero, i. e. when the longitude difference $l - l_c$ of the direction in question from that of the centre of the Galaxy is close to $\frac{1}{2}\, n\pi$, where n is an integer.

The further study of the individual components of the lines, and of their intensities and contours, should lead to a discovery of the physical properties of the separate clouds of interstellar gas and of their chemical composition. The above conclusions concerning the presence of a system of numerous clouds of interstellar gas in the Galaxy lead, in turn, to the problem of the extent to which this system coincides or is identical with that of the numerous clouds of interstellar dust which cause the general absorption.

The use of nebular spectrographs has made possible the conclusion that there exist in the Galaxy, besides the diffuse gaseous nebulae

(like the Orion nebula, and the Rosette nebula around the cluster NGC 2244 in Monoceros), large faintly luminous *hydrogen field nebulae* (so called because they are detected mainly by their radiation in the lines of the Balmer series). Both the diffuse gaseous nebulae and these hydrogen fields are *gaseous clouds with emission.* On the other hand, there is no possibility of observing interstellar hydrogen absorption lines in the optical region of the spectrum, since all the neutral hydrogen atoms in the gaseous interstellar clouds are in the ground state, as a result of which these clouds can absorb only lines of the Lyman series, which are in the region of the spectrum absorbed by the Earth's atmosphere.

Hence the problem of the relation between the interstellar gas clouds, the diffuse nebulae and the hydrogen fields is rather complex. The question arises to what extent the diffuse nebulae and the faint hydrogen fields can be regarded as particular cases of interstellar gas clouds, differing from other clouds only by the presence of conditions under which emission lines appear (i. e. the presence of short-wave radiation from hot stars).

The impossibility of a direct investigation of the distribution of the immense mass of hydrogen in interstellar space, since almost all its atoms are in the ground state and the excitation is very slight in the majority of the gaseous clouds, until recently caused a considerable indeterminacy in our conclusions on the properties of these clouds. However, the discovery in 1951 of radio waves with a frequency of 1420·4 Mc/s, corresponding to the transition between two sub-levels of the ground level $^2S_{\frac{1}{2}}$ of the neutral hydrogen atom (the sub-levels are due to the hyperfine structure of the hydrogen levels, which is related to the existence of the magnetic moment of the proton), has made possible a direct investigation of hydrogen in the normal non-excited state, and should lead in the near future to a considerable increase in our knowledge of the distribution and motion of the gaseous interstellar matter.

Recent investigations of the radiation at a frequency of 1420 Mc/s (the 21 cm line) have not only confirmed our expectations, but also provided the solution of such problems as, for instance, the determination of the position of the spiral arms of the Galaxy (by the work of J. H. Oort and others). The possibility has been noticed of finding exactly the radial velocities of separate interstellar clouds and so identifying the clouds observed by radio astronomy with those found from the interstellar sodium and calcium lines in the visible part of the spectrum.

For further investigations in the field, and to resolve the problems stated above, it will be necessary to extend the theory of the formation of interstellar absorption and emission lines and that describing the state of the interstellar gas.

2. The formation of interstellar absorption lines. The theory of the formation of absorption lines when the light of a star passes through some cloud of interstellar gas is simpler than that of the absorption lines formed in the atmospheres of the stars themselves. The reason for this is that we can neglect the scattered light in the frequencies of the absorption lines of the interstellar gas. For, although the light of a star is scattered in all directions on being absorbed by a nebula in the frequencies of resonance lines (for instance H and K), only a negligible part of it is scattered in the direction of the original radiation from the star. Hence we observe almost entirely a simple attenuation (*extinction*) of the light of the star in the line frequencies.

In this case, the intensity of the radiation after passing through the cloud is determined by the formula

$$I_\lambda = I_\lambda^0 \, e^{-\tau_\lambda},$$

where τ_λ is the optical thickness of the cloud in wavelength λ, and I_λ^0 is the intensity of the star's light before attenuation. Hence we have for the equivalent width of the line

$$W = \int (1 - e^{-\tau_\lambda}) \, d\lambda \,. \tag{36.1}$$

The optical thickness τ_λ can be expressed in terms of the number N of absorbing atoms in a cylinder of cross-section 1 cm^2, whose axis coincides with the line of sight and extends through the cloud:

$$\tau_\lambda = N s_\lambda \,, \tag{36.2}$$

where s_λ is the mean absorption coefficient referred to one atom, or, more precisely, the mean absorption coefficient when there is one atom per unit volume.

The dependence of s_λ on λ has been shown in Part II to be determined by the various causes of the broadening of a spectral line. Under the conditions existing in the rarefied interstellar gas, only two causes of line broadening can be at all important: (1) the broadening by radiation damping and (2) the Doppler effect, both thermal and macroscopic. No kind of broadening by collisions or by the molecular Stark effect can play any part. The natural broadening by radiation damping is small in comparison with the thermal Doppler broadening. It can begin to be important only in the case where the number of absorbing atoms is so large that the medium becomes opaque even in the wings of the absorption lines beyond the Doppler width. However, the observed interstellar lines, and particularly the components of these lines which belong to individual clouds, are so narrow that we cannot speak of any absorption in the wings of the lines. This points to a small number N of absorbing atoms along the path of a ray.

Thus only one important factor remains to be taken into account, namely the thermal Doppler broadening. Under these conditions, the dependence of the absorption coefficient on the wavelength should be expressed by the formula

$$s_\lambda = s_0 \, e^{-c^2(\lambda - \lambda_0)^2/\lambda_0^2 b^2} , \tag{36.3}$$

where c is the velocity of light, b is given in terms of the temperature and the atomic mass by the formula

$$b = \sqrt{(2\, kT/m)} , \tag{36.4}$$

and λ_0 is the wavelength of the line which corresponds to the velocity of the cloud concerned.

The constant s_0, which is equal to the absorption coefficient at the centre of the line, can be found by starting from the fact that the total amount of energy (expressed in terms of s_λ) absorbed in unit time in the line, when there is isotropic radiation of constant intensity inside the line, is equal to the same amount absorbed in unit time, expressed in terms of the Einstein transition probability, i. e.

$$4\,\pi\, I_\nu \int s_\lambda \,\mathrm{d}\lambda = B_{1\to 2}\, \varrho_\nu\, h\nu ,$$

where ϱ_ν is the density of radiation. From this relation, taking into account that

$$\varrho_\nu = 4\,\pi\, I_\nu/c \text{ and } B_{1\to 2} = (g_2/g_1)\, A_{2\to 1}\, c^3/8\,\pi\, h\, \nu^3 ,$$

we find

$$\int s_\lambda \,\mathrm{d}\lambda = (g_2/g_1)\, A_{2\to 1}\, \lambda^2/8\,\pi , \tag{36.5}$$

where g_1 and g_2 are the statistical weights of the lower and upper states. Substituting the expression (36.3) in (36.5), we obtain, after integration,

$$s_0 = (g_2/g_1)\, A_{2\to 1}\, \lambda\, c/8\, \pi^{3/2}\, b . \tag{36.6}$$

On the other hand, substituting (36.3) in (36.2) and thence in (36.1), we find

$$W = \int (1 - e^{-Ns_0 \exp(-c^2(\lambda-\lambda_0)^2/\lambda_0^2 b^2)})\,\mathrm{d}\lambda , \tag{36.7}$$

or, putting

$$c(\lambda - \lambda_0)/\lambda_0 b = x ,$$

we have

$$W = \frac{\lambda_0 b}{c} \int_{-\infty}^{\infty} (1 - e^{-Ns_0 \exp(-x^2)})\,\mathrm{d}x . \tag{36.8}$$

The definite integral on the right-hand side is some function of Ns_0, which can be tabulated. We denote this function by $F(Ns_0)$. Thus

$$W/b = \lambda_0 F(Ns_0)/c\,. \tag{36.9}$$

The argument Ns_0 of the function F is the optical thickness τ_0 at the centre of the line. Table 36 gives the values of $\log_{10}\tau_0$ and the corresponding values of $\log_{10}F(\tau_0)$.

Table 36

$\log_{10}\tau_0$	$\log_{10}F(\tau_0)$	$\log_{10}\tau_0$	$\log_{10}F(\tau_0)$
$\bar{1}\cdot0$	$\bar{1}\cdot24$	0·6	0·42
$\bar{1}\cdot2$	$\bar{1}\cdot43$	0·8	0·48
$\bar{1}\cdot4$	$\bar{1}\cdot61$	1·0	0·52
$\bar{1}\cdot6$	$\bar{1}\cdot79$	1·2	0·56
$\bar{1}\cdot8$	$\bar{1}\cdot96$	1·4	0·59
0·0	0·11	1·6	0·62
0·2	0·24	1·8	0·64
0·4	0·34	2·0	0·66

It is seen from this table that, for small values of τ_0, the equivalent width is proportional to τ_0, but for $\tau_0 > 1$ saturation begins and the equivalent width increases more and more slowly.

Observations gives the value W of the equivalent width of a spectral line. The transition probability $A_{2\to1}$ which appears in s_0, according to (36.6), should be known from physics. However, two unknown quantities N and b characterising the cloud appear in formula (36.9). Moreover, as we see from (36.6), b appears also in the denominator of the expression for s_0. In order to determine N and b simultaneously, it is necessary to know the equivalent widths of at least two lines absorbed by the ground state of the atom concerned. We shall show how this may be done.

If the lines correspond to two different transitions, say $1 \to k$ and $1 \to l$, we have for the ratio of the equivalent widths W_{1k} and W_{1l}, from (36.9),

$$\frac{\lambda_{1l}}{\lambda_{1k}}\,\frac{W_{1k}}{W_{1l}} = \frac{F(N\,s_{0,1k})}{F(N\,s_{0,1l})}\,. \tag{36.10}$$

The left-hand side of (36.10) can be found from observation. We can therefore find the difference

$$\log_{10}F(N\,s_{0,1k}) - \log_{10}F(N\,s_{0,1l}) = \log_{10}(\lambda_{1l}\,W_{1k}/\lambda_{1k}\,W_{1l})\,. \tag{36.11}$$

On the other hand, from (36.6) the ratio

$$\frac{\tau_{0,1k}}{\tau_{0,1l}} = \frac{s_{0,1k}}{s_{0,1l}} = \frac{g_k}{g_l}\,\frac{A_{k\to1}}{A_{l\to1}}\,\frac{\lambda_{1k}}{\lambda_{1l}} \tag{36.12}$$

can be calculated from physical data. Thus the difference

$$\Delta = \log_{10} \tau_{0,1k} - \log_{10} \tau_{0,1l} \tag{36.13}$$

can also be calculated.

Having calculated (36.11) and (36.13), we can find in Table 36 the pair of values of $\log_{10} \tau_0$ which have the given difference (36.13) and lead to the difference in $\log_{10} F$ calculated by (36.11).

Thus $\tau_{0,1k}$, $\tau_{0,1l}$, $F(N s_{0,1k})$ and $F(N s_{0,1l})$ are determined. Knowing $F(N s_{0,1k})$, and W_{1k}, from (36.9) we at once find b, and hence we obtain N from the relation $\tau_{0,1k} = N s_{0,1k}$, using (36.6).

As an example, let us take interstellar sodium. We might attempt, in this case, to use the ratio of the components of the doublet D_1 and D_2. The ratio of the values of s_0 is 2. However, for very large W the ratio of the equivalent widths rapidly approaches unity, as is seen from Table 36. Hence a slight error in the empirical determination of the ratio of equivalent widths of the components can result in a large error in the determination of the values of τ_0 for these components, and consequently a large error in the determination of N.

However, we can also use the ratio of equivalent widths for the D_1 line in the yellow doublet and, say, the line 3303 Å in the ultra-violet doublet. In this case the ratio of the transition probabilities is known also. It has been determined by Yu. V. Filippov and V. K. Prokof'ev and is 21·8. For such a large ratio of transition probabilities, and therefore of the optical depths τ_0 in the centres of the two lines, the values of τ_0 are determined much more accurately from the known differences of $\log_{10} \tau_0$ and $\log_{10} F(\tau_0)$. The value of N is therefore determined more exactly. Thus, from the ratio $D_1/3303$ observed for the star χ^2 Orionis, which is 9·8, B. Strömgren has found that $\log_{10} N = 13{\cdot}99$, i. e. the number of absorbing sodium atoms is $9{\cdot}8 \times 10^{13}$ per cm^2. By carrying out a similar analysis for the lines of various atoms and ions, we can draw corresponding conclusions about their relative abundance.

The quantity b has been introduced in connection with the thermal motion of the absorbing atoms. However, it may happen that there are also macroscopic motions in the cloud. The velocity dispersion of these motions along the line of sight results in an increase in the effective value of b. Similarly, in the case where the components of the same line, corresponding to different clouds which lie in the line of sight, coalesce (which usually occurs in directions where the galactic rotation has only a small effect on the radial velocities), the velocity dispersion of the gaseous clouds through which the star's light passes has an effect on b.

A comparison of the values of N for Na and Ca II, obtained by the above method, shows that these quantities are often of the same order for the same star. Thus, for example, according to L. SPITZER, the main component of the interstellar lines in the spectrum of ξ Persei, i. e. the component due to the denser cloud, gives for Na I the value $N = 4 \times 10^{12}$ per cm^2, and for Ca II $1{\cdot}7 \times 10^{12}$ per cm^2. The value of N for Na I is sometimes as much as four times that for Ca II. These facts are explained both by differences in the abundances of the two elements in the interstellar cloud and by different degrees of ionisation.

3. **The ionisation of the interstellar gas.** Apart from those cases where a gaseous cloud lies immediately next to a hot supergiant or envelops it, i. e. the cases where we are concerned with conditions similar to those existing in the bright diffuse nebulae, the ionisation of the matter in the interstellar gas clouds is due to the simultaneous effect of the radiation from all the stars in the Galaxy. Here it is evident that we have, firstly, greatly diluted radiation from the stars and, secondly, a superposition of radiations corresponding to the various temperatures of the illuminating stars. The ionisation by electron collision plays only a small part under the conditions existing in the interstellar gas.

Let us examine the ionisation at some point in an interstellar cloud. We denote by $\mathrm{d}\Omega$ the total solid angle subtended at this point by all stars having temperatures between T and $T + \mathrm{d}T$. Assuming that Planck's Law holds for these stars, we find that the density of radiation ϱ_ν due to the stars of this one temperature group is

$$\varrho_\nu = \frac{2\,h\,\nu^3}{c^3}\,\frac{\mathrm{d}\Omega}{e^{h\nu/kT} - 1}\,.$$

Hence the total density of radiation will be

$$\varrho_\nu = \frac{2\,h\,\nu^3}{c^3}\int\limits_0^\infty \frac{1}{e^{h\nu/kT} - 1}\,\frac{\mathrm{d}\Omega}{\mathrm{d}T}\,\mathrm{d}T\,. \tag{36.14}$$

Thus, to calculate ϱ_ν, it is necessary to know the function $\mathrm{d}\Omega/\mathrm{d}T$. In the present state of our knowledge, we can determine this function only for the neighbourhood of the Sun, using the known data on the apparent brightness of the stars and their distribution among the spectral classes. The relative frequency distribution of density of radiation thus obtained is markedly different from the Planckian. Hence the ionisation of the interstellar gas cannot be expressed by the formulae which hold when Planckian radiation is present, nor (as is done in the case of the planetary nebulae, where we have dilute Planckian radiation) by introducing a factor W into the formula for the radiation

density. The calculation of the degree of ionisation for various kinds of atoms or ions therefore requires fairly detailed numerical calculations, based on the numerical values of ϱ_ν and on the constancy of the numbers of atoms in various states of ionisation and excitation.

Although various rather unjustified assumptions have to be made, in these calculations, concerning the number of free electrons, we reach in every case the conclusion that the sodium atoms must be much more highly ionised than the calcium ions. We find for the ratio of the numbers of sodium ions and neutral sodium atoms

$$n_1/n \sim 10^6\,, \tag{36.15}$$

and for the ratio of the numbers of doubly and singly ionised calcium atoms

$$n_2/n_1 \sim 10^4\,. \tag{36.16}$$

In consequence of the very high ionisation potentials of Na II and Ca III, the second ionisation of sodium and the third ionisation of calcium can occur only extremely rarely. Hence [see formulae (36.15) and (36.16)] the overwhelming majority of sodium atoms must be in the singly ionised state, and of calcium atoms in the doubly ionised state.

If we further assume that the total number of sodium atoms in the interstellar gas, in all stages of ionisation, is three times the total number of calcium atoms (as is true for the Sun's atmosphere), the number of Ca II atoms must be approximately 30 times the number of neutral sodium atoms. The lines of the H and K doublet, however, have equivalent widths of the same order as those of the $D_{1,2}$ yellow doublet lines. An attempt has been made to explain this by the phenomenon of saturation of absorption lines as the number of atoms increases. The saturation phenomenon is theoretically accounted for by the results in Table 36.

However, equivalent widths, and numbers of atoms, of the same order are often obtained for the Na I and Ca II doublets even when these and other lines are far from being saturated. It is therefore possible that sodium occurs much more abundantly than calcium in the gaseous interstellar matter, its relative abundance being greater than in the Sun.

The absence of data on the hydrogen absorption lines of interstellar matter greatly hinders the study of the state of the gas in the interstellar clouds. However, observations have established the presence of interstellar clouds which emit lines of the Balmer series, the clouds being most easily observed in the H_α line.

4. The H_α radiation fields. We have seen above that the number of neutral sodium atoms along the line of sight in a single cloud sometimes

reaches 10^{13} or 10^{14} per cm^2. Since hydrogen is a thousand times more abundant than sodium, and the degree of ionisation of hydrogen should be at least two orders of magnitude less, we reach the conclusion that the number of neutral hydrogen atoms in a cylinder of base 1 cm^2 along the line of sight should amount to 10^{17}, 10^{18} or more. This means that the gaseous interstellar clouds should be opaque to the continuous radiation beyond the limit of the Lyman series. This result leads to interesting consequences. Let us imagine a star of class O or B 0, or a group of such stars, lying close to the galactic plane. If such a star (or group of stars) lies inside one of the clouds, then it should cause a high ionisation within this cloud and a bright emission of the hydrogen lines. In this case we should observe a luminous diffuse gaseous nebula. If, however, the star is between the gaseous clouds, it should cause ionisation in the clouds nearest to it. We should observe the radiation of these clouds in the lines of the Balmer series, i. e. we should see, around the exciting star, extensive fields of H_α radiation. These have been detected by means of nebular spectrographs.

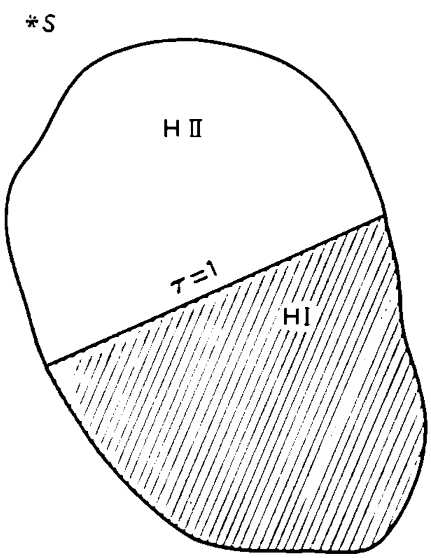

FIG. 75

It is important to note, however, that the clouds illuminated by the exciting star must completely absorb the radiation beyond the limit of the Lyman series, as a result of which the other more distant clouds screened by them do not receive any L_c radiation at all, and hence no hydrogen radiation is excited in them.

This happens as follows (see Fig. 75). The L_c quanta emitted by the star S penetrate into the cloud to an optical depth $\tau = 1$ in the L_c frequencies. Beyond this, however, the radiation is attenuated, and so a marked decrease is caused in the degree of ionisation; consequently, the number of hydrogen atoms in the neutral state rapidly increases. Hence, from the layer $\tau = 1$ onwards, the optical depth increases much more rapidly than it does up to $\tau = 1$. As a result, the absorption of L_c quanta in these layers is increased. These facts mean that the part of the cloud lying beyond the layer $\tau = 1$ must have a very great optical thickness in the L_c frequencies. The cloud is composed of an ionised part (the H II region) and an un-ionised part (the H I region). Consequently the clouds lying at a great distance from the star, behind the cloud we are considering, can hardly be excited at all by the radiation of the star in the hydrogen continuum.

In the same way as the masses of the planetary and diffuse gaseous nebulae are determined, we can determine the mass of hydrogen that is concentrated in the luminous part of an interstellar cloud. Here we must start from the condition that the radiation from unit volume in each Balmer line is proportional to the number of recombinations of hydrogen atoms:

$$\varepsilon = C\, n_i\, n_e = C\, n_i^2 \,.$$

Multiplying ε by the square of the volume V, we have

$$\varepsilon\, V^2 = C\, n_i^2\, V^2 = C\, \mathsf{M}^2/m^2 \,,$$

where m is the mass of a hydrogen atom, and M is the mass of *the luminous part* of the nebula. The quantity $\varepsilon\, V$ is equal to the luminosity L of the nebula in the line of the Balmer series considered. Hence

$$L\, V = C\, \mathsf{M}^2/m^2 \,. \tag{36.17}$$

Formula (36.17) allows M to be determined from L and V.

The masses of the interstellar clouds obtained from such a calculation are of the order of ten times the Sun's mass. The number of all hydrogen atoms, both ionised and neutral, in a cubic centimetre is of the order of ten. The radii of the clouds are found from observation to be of the order of ten parsecs.

5. **Giant gaseous nebulae.** At the Crimean Astrophysical Observatory, Academician G. A. Shaĭn has developed a new method of detecting and investigating bright diffuse gaseous nebulae. This method has enabled G. A. Shaĭn and V. F. Gaze [145] to discover a considerable number of new gaseous nebulae and to reach a deeper understanding of their nature. B. J. Bok and his co-workers have discovered a number of such nebulae in the southern hemisphere. In particular, the determination by G. A. Shaĭn and V. F. Gaze of the masses of the diffuse nebulae, using estimates of their luminosity in the lines of the Balmer series, has led to the recognition of the group of *giant gaseous nebulae*, having masses of the order of hundreds or thousands of times the Sun's mass, or even more. Of the well-known diffuse nebulae, this group contains the Orion nebula and the Rosette nebula around the O cluster NGC 2244. It has been found that the cluster NGC 1805, which is the nucleus of the O association Cassiopeia VI, is also surrounded by uch a giant diffuse nebula. An interesting giant gaseous nebula has been discovered in the association Cassiopeia II, not far from the nucleus, which is the cluster NGC 7510.

It is interesting that all these giant nebulae lie in the neighbourhood of O associations. The actual location of these nebulae, relative to the stars in the association which cause their radiation, indicates a profound genetic relation between the nebulae and the members of the corresponding associations. These giant diffuse nebulae can scarcely be supposed to be ordinary representatives of the interstellar gas clouds, which have by chance come into the vicinity of members of associations of hot giants. Hence the conclusion that the diffuse dust nebulae are casually associated with the stars which illuminate them cannot be extended to these giant gaseous nebulae.

The determination of the masses of some giant diffuse nebulae lying in the nearest extra-galactic systems has led to values of the order of ten thousand times the Sun's mass. The situation of these supergiant diffuse nebulae coincides with that of associations of blue giants found in these systems.

Another important fact discovered in recent investigations of these nebulae is that a large number of filamentary nebulae exist in the Galaxy. This has led in turn to a more detailed study of the individual filaments in diffuse nebulae known previously. G. A. Shaĭn has compared the filamentary structure of the gas-and-dust nebulae with the drawn-out form of many dark nebulae, which are sometimes seen as dark lanes against the background of the star field. It was natural to suppose that the directions of the filaments and lanes are those of the magnetic lines of force in the region of space considered. This hypothesis can be tested by studying the polarisation of the light from stars near these objects. Such a test has been made by G. A. Shaĭn, and the result was favourable. The result of these investigations is to reveal the importance of interstellar magnetic fields in the motion of diffuse matter. Since the interstellar gas is strongly ionised, this importance is easily understood from the point of view of magnetohydrodynamics.

6. **The Crab nebula.** Among the nebulae observed in the Galaxy, the Crab nebula is unique in appearance. The continuous spectrum emitted by the central part of this nebula has always been a source of interest. The reason for this is that the continuous spectrum cannot be regarded as a reflection of the light of a star, since there is no bright illuminating star in this case. It was clear that the source of the radiation is in the nebula itself, yet it was difficult to give an account of the physical nature of this source.

Some time after it had been discovered that the Crab nebula is one of the most intense discrete radio sources, I. S. Shklovskiĭ put forward the hypothesis that the continuous spectrum in optical frequencies is essentially a continuation of the observed radio spectrum. Similarly to the discussion in Chapter 35, I. S. Shklovskiĭ deduced that this con-

tinuous spectrum is the radiation of relativistic electrons in a magnetic field. It was found that, to explain the optical continuous spectrum, we must assume that the energies of the relativistic electrons reach values of 10^{11} eV or higher. I. M. GORDON, of the Kharkov observatory, put forward the idea that this hypothesis might be tested by observations of the polarisation of the Crab nebula. In 1953—1954, V. A. DOMBROVSKIĬ, using observations made at the Byurakan observatory, showed that the light of the Crab nebula is polarised to an extraordinarily high degree (up to 20 %). According to further observations made at Byurakan, the degree of polarisation in particular regions of the nebula reaches 60 %. For comparison, it may be mentioned that the greatest polarisation of starlight is 6 %.

One of the most important problems of present-day astrophysics is to discover the sources of the high-energy electrons in the Crab nebula. It has recently been pointed out that the continuous spectrum of certain cometary nebulae also cannot be explained entirely by reflection of light from a star. In this case, too, we are concerned with a continuous emission of non-thermal character.

Appendix I

Tables for calculating the selective absorption coefficient when Doppler broadening and radiation damping act simultaneously

The following expression for the function $\phi(a,p) = s_\nu/s_{\nu_0}$, defined by (11.39), can be obtained* for values of $a = \delta_{ik}/\Delta\nu_D$ which satisfy the inequality $a \ll 1$:

$$\phi(a,p) = s_\nu/s_{\nu_0} = e^{-p^2} - \frac{2a}{\sqrt{\pi}}\left\{1 - 2\,p\,e^{-p^2}\int_0^p e^{z^2}\,\mathrm{d}z\right\}. \qquad \text{(A.1)}$$

Table A.1 gives, for various values of p, the values of the quantity in braces on the right-hand side of this formula,

$$\Phi(p) = 1 - 2\,p\,e^{-p^2}\int_0^p e^{z^2}\,\mathrm{d}z\,. \qquad \text{(A.2)}$$

For $a = 0{\cdot}03$, the error in the values of $\phi(a,p)$ arising from the use of formula (A.1) and Table A.1 does not exceed 0·5 % of the value of $\phi(a,p)$ itself. For $a = 0{\cdot}10$ the maximum error is about 3 %, and for $a = 0{\cdot}20$ about 9 %. Thus, for values of a larger than about 0·05 it is desirable to have tables giving greater accuracy than Table A.1. We give here three such tables, compiled by F. HJERTING [61]. Table A.2 gives the function $\phi(a,p) = s_\nu/s_{\nu_0}$ for a series of values of a and p; the greatest value of a in the table is 0·2, and the greatest value of p is 5. Table A.3 gives the values of $s_\nu/a s_{\nu_0}$ for p varying from 5 to 20.

Finally, Table A.4 gives the values of s_ν/s_{ν_0} for a varying from 0·0 to 0·5.

There is also a useful table compiled by D. L. HARRIS [57].

* See, e. g., A. UNSÖLD [168, p. 263]. It must be borne in mind that the quantity a appearing in UNSÖLD's formula is twice that defined by our formula (11.38).

Table A.1

p	$\Phi(p)$	p	$\Phi(p)$	p	$\Phi(p)$
0·0	1·0000	4·2	—0·03119	8·4	—0·007242
0·2	0·9221	4·4	—0·02815	8·6	—0·006902
0·4	0·7121	4·6	—0·02554	8·8	—0·006586
0·6	0·4303	4·8	—0·02330	9·0	—0·006290
0·8	0·1487	5·0	—0·02134	9·2	—0·006014
1·0	—0·07616	5·2	—0·01963	9·4	—0·005757
1·2	—0·2175	5·4	—0·01812	9·6	—0·005516
1·4	—0·2782	5·6	—0·01678	9·8	—0·005290
1·6	—0·2797	5·8	—0·01558	10·0	—0·005076
1·8	—0·2485	6·0	—0·01451	10·2	—0·004877
2·0	—0·2054	6·2	—0·01355	10·4	—0·004688
2·2	—0·1638	6·4	—0·01268	10·6	—0·004511
2·4	—0·1295	6·6	—0·01190	10·8	—0·004344
2·6	—0·1033	6·8	—0·01118	11·0	—0·004183
2·8	—0·08389	7·0	—0·01053	11·2	—0·004035
3·0	—0·06962	7·2	—0·009938	11·4	—0·003893
3·2	—0·05896	7·4	—0·009393	11·6	—0·003757
3·4	—0·05076	7·6	—0·008892	11·8	—0·003630
3·6	—0·04430	7·8	—0·008429	12·0	—0·003510
3·8	—0·03908	8·0	—0·008003		
4·0	—0·03478	8·2	—0·007608		

Table A.2

a	p = 0·00	0·25	0·50	0·75	1·00
0·00	1·000	0·939	0·779	0·570	0·368
0·01	0·989	0·930	0·772	0·567	0·369
0·02	0·978	0·920	0·766	0·565	0·370
0·03	0·967	0·910	0·760	0·562	0·370
0·04	0·956	0·901	0·753	0·560	0·371
0·05	0·946	0·892	0·747	0·558	0·371
0·06	0·936	0·883	0·741	0·555	0·372
0·07	0·926	0·874	0·735	0·552	0·372
0·08	0·916	0·865	0·729	0·550	0·373
0·09	0·906	0·856	0·723	0·548	0·373
0·10	0·896	0·848	0·718	0·545	0·373
0·11	0·887	0·840	0·712	0·542	0·374
0·12	0·878	0·831	0·706	0·540	0·374
0·13	0·869	0·823	0·701	0·538	0·374
0·14	0·860	0·815	0·695	0·535	0·374
0·15	0·851	0·807	0·690	0·532	0·374
0·16	0·842	0·799	0·684	0·530	0·374
0·17	0·834	0·792	0·679	0·527	0·374
0·18	0·825	0·784	0·674	0·525	0·374
0·19	0·817	0·777	0·668	0·522	0·373
0·20	0·809	0·770	0·663	0·520	0·373

a	p = 2·75	3·00	3·25	3·50	3·75
0·00	0·00052	0·000123	0·000026	0·000005	0·000001
0·01	0·00151	0·000908	0·000666	0·000539	0·000455
0·02	0·00251	0·00169	0·00130	0·00107	0·000910
0·03	0·00350	0·00248	0·00194	0·00161	0·00136
0·04	0·00449	0·00326	0·00258	0·00214	0·00182
0·05	0·00548	0·00404	0·00322	0·00268	0·00227
0·06	0·00646	0·00483	0·00386	0·00321	0·00273
0·07	0·00745	0·00561	0·00450	0·00374	0·00318
0·08	0·00843	0·00639	0·00513	0·00428	0·00364
0·09	0·00940	0·00717	0·00577	0·00481	0·00409
0·10	0·0104	0·00794	0·00641	0·00534	0·00454
0·11	0·0114	0·00872	0·00704	0·00587	0·00500
0·12	0·0123	0·00949	0·00768	0·00640	0·00545
0·13	0·0133	0·0103	0·00831	0·00693	0·00590
0·14	0·0142	0·0110	0·00894	0·00746	0·00635
0·15	0·0152	0·0118	0·00957	0·00799	0·00680
0·16	0·0162	0·0126	0·0102	0·00852	0·00726
0·17	0·0171	0·0133	0·0108	0·00905	0·00771
0·18	0·0180	0·0141	0·0115	0·00958	0·00816
0·19	0·0190	0·0149	0·0121	0·0101	0·00860
0·20	0·0199	0·0156	0·0127	0·0106	0·00906

Table A.2

p

1·25	1·50	1·75	2·00	2·25	2·50
0·210	0·105	0·0468	0·0183	0·00633	0·00193
0·212	0·109	0·0497	0·0206	0·00806	0·00323
0·215	0·112	0·0525	0·0229	0·00979	0·00452
0·217	0·115	0·0553	0·0252	0·0115	0·00581
0·220	0·118	0·0581	0·0274	0·0132	0·00710
0·222	0·121	0·0608	0·0296	0·0149	0·00838
0·224	0·123	0·0635	0·0318	0·0166	0·00966
0·226	0·126	0·0660	0·0339	0·0182	0·0109
0·229	0·129	0·0687	0·0360	0·0198	0·0122
0·231	0·131	0·0713	0·0381	0·0215	0·0134
0·232	0·134	0·0737	0·0402	0·0231	0·0147
0·234	0·136	0·0762	0·0422	0·0247	0·0160
0·236	0·139	0·0786	0·0442	0·0263	0·0172
0·238	0·141	0·0810	0·0463	0·0279	0·0184
0·240	0·144	0·0833	0·0482	0·0294	0·0196
0·241	0·146	0·0856	0·0502	0·0310	0·0208
0·243	0·148	0·0879	0·0520	0·0325	0·0221
0·244	0·150	0·0901	0·0540	0·0340	0·0233
0·246	0·152	0·0922	0·0558	0·0356	0·0245
0·247	0·154	0·0943	0·0576	0·0371	0·0256
0·248	0·157	0·0964	0·0595	0·0385	0·0268

p

4·00	4·25	4·50	4·75	5·00
0·000000	0·000000	0·000000	0·000000	0·000000
0·000393	0·000343	0·000302	0·000269	0·000241
0·000785	0·000686	0·000605	0·000538	0·000482
0·00118	0·00103	0·000907	0·000806	0·000722
0·00157	0·00137	0·00121	0·00108	0·000963
0·00196	0·00171	0·00151	0·00134	0·00120
0·00236	0·00206	0·00181	0·00161	0·00144
0·00275	0·00240	0·00212	0·00188	0·00169
0·00314	0·00274	0·00242	0·00215	0·00192
0·00353	0·00308	0·00272	0·00242	0·00217
0·00392	0·00343	0·00302	0·00269	0·00241
0·00432	0·00377	0·00332	0·00295	0·00265
0·00471	0·00411	0·00362	0·00322	0·00289
0·00510	0·00445	0·00392	0·00349	0·00313
0·00549	0·00479	0·00423	0·00376	0·00337
0·00588	0·00513	0·00453	0·00403	0·00361
0·00627	0·00547	0 00483	0·00429	0·00385
0·00666	0·00581	0·00513	0·00456	0·00409
0·00705	0·00615	0·00543	0·00483	0·00433
0·00744	0·00649	0·00573	0·00510	0·00457
0·00783	0·00683	0·00603	0·00536	0·00481

Table A.3

p	$\frac{s_\nu}{s_{\nu_0}} a^{-1}$	p	$\frac{s_\nu}{s_{\nu_0}} a^{-1}$	p	$\frac{s_\nu}{s_{\nu_0}} a^{-1}$
5·0	0·0241	10·0	0·00573	15·0	0·00251
5·5	0·0197	10·5	0·00519	15·5	0·00235
6·0	0·0164	11·0	0·00472	16·0	0·00221
6·5	0·0139	11·5	0·00432	16·5	0·00208
7·0	0·0119	12·0	0·00396	17·0	0·00196
7·5	0·0103	12·5	0·00365	17·5	0·00185
8·0	0·00903	13·0	0·00337	18·0	0·00175
8·5	0·00798	13·5	0·00312	18·5	0·00166
9·0	0·00710	14·0	0·00290	19·0	0·00157
9·5	0·00636	14·5	0·00270	19·5	0·00149
10·0	0·00573	15·0	0·00251	20·0	0·00142

Table A.4

p	s_ν/s_{ν_0}					
	$a = 0{\cdot}0$	$a = 0{\cdot}1$	$a = 0{\cdot}2$	$a = 0{\cdot}3$	$a = 0{\cdot}4$	$a = 0{\cdot}5$
0·00	1·00	0·90	0·81	0·74	0·67	0·62
0·25	0·94	0·85	0·77	0·70	0·64	0·59
0·50	0·78	0·72	0·66	0·62	0·57	0·53
0·75	0·57	0·54	0·52	0·50	0·47	0·44
1·00	0·37	0·37	0·37	0·37	0·36	0·36
1·25	0·21	0·23	0·25	0·26	0·27	0·27
1·50	0·10	0·13	0·16	0·17	0·19	0·20
1·75	0·047	0·074	0·096	0·12	0·13	0·14
2·00	0·018	0·040	0·060	0·076	0·090	0·103
2·50	0·002	0·015	0·027	0·038	0·049	0·058
3·00	0·0001	0·0079	0·016	0·023	0·030	0·037
3·50	0·0000	0·0053	0·011	0·016	0·021	0·026
4·00	0·0000	0·0039	0·0078	0·012	0·016	0·019
4·50	0·0000	0·0030	0·0060	0·0090	0·012	0·015
5·00	0·0000	0·0024	0·0048	0·0072	0·0096	0·012

Appendix II

Some of the most important astronomical and physical constants

1. Astronomical constants

One astronomical unit	$1{\cdot}4964 \times 10^{13}$ cm
One light year	$9{\cdot}463 \times 10^{17}$ cm
One parsec = 3·26 light years	$3{\cdot}084 \times 10^{18}$ cm
Mass of the Sun	$1{\cdot}991 \times 10^{33}$ g
Radius of the Sun	$6{\cdot}963 \times 10^{10}$ cm
Mean density of the Sun	$1{\cdot}41$ g/cm^3
Acceleration due to gravity at the Sun's surface	$2{\cdot}740 \times 10^{4}$ cm/sec^2
Total energy emitted by the Sun	$3{\cdot}79 \times 10^{33}$ erg/sec
Number of seconds in one year	$3{\cdot}1558 \times 10^{7}$ sec
Sun's parallax	$8''{\cdot}790$

2. Physical constants

Velocity of light in vacuo	c	$= 2{\cdot}99776 \times 10^{10}$ cm/sec
Gravitational constant	G	$= 6{\cdot}670 \times 10^{-8}$ dyn cm^2/g^2
Planck's constant	h	$= 6{\cdot}624 \times 10^{-27}$ erg sec
Avogadro's constant (number of atoms or molecules in one mole)	N_0	$= 6{\cdot}023 \times 10^{23}$ per mole
Charge on the electron	e	$= 4{\cdot}802 \times 10^{-10}$ CGS units
Rydberg's constant for hydrogen	R_H	$= 109{,}677{\cdot}581$ cm^{-1}
Mass of the electron	m_e	$= 9{\cdot}1066 \times 10^{-28}$ g
Mass of a hydrogen (^{1}H) atom	m_H	$= 1{\cdot}6734 \times 10^{-24}$ g
Mass of the proton	m_p	$= 1{\cdot}6725 \times 10^{-24}$ g
Ratio of mass of the proton to mass of the electron	m_p/m_e	$= 1836{\cdot}5$
Boltzmann's constant	k	$= 1{\cdot}380 \times 10^{-16}$ erg/deg
First constant in the radiation law	c_1	$= 8\pi hc = 4{\cdot}990 \times 10^{-15}$ erg cm
Second constant in the radiation law	c_2	$= hc/k = 1{\cdot}4384$ cm deg
Constant in the Stefan-Boltzmann law	σ	$= 5{\cdot}672 \times 10^{-5}$ erg/cm^2 deg^4 sec
Displacement constant in Wien's law	$\lambda_{max}T$	$= 0{\cdot}2897$ cm deg
Radius of the first orbit in the hydrogen atom	a_0	$= h^2/4\pi^2 m_e e^2 = 0{\cdot}529 \times 10^{-8}$ cm
Number of ergs in one electron-volt		$1{\cdot}6020 \times 10^{-12}$ erg
Wavelength of a quantum of energy 1 eV		12,395 Å
Temperature corresponding to 1 eV		11,606° K
Standard atmosphere		1,013,246 dyn/cm^2 atm
Mechanical equivalent of heat		4·1855 J/cal
Gas constant per mole	R	$= 8{\cdot}3144 \times 10^{7}$ erg/deg mole
Radiation density constant	a	$= 7{\cdot}569 \times 10^{-15}$ erg/cm^3 deg^4

Appendix III

D. I. Mendeleev's Periodic System of the Elements

Groups of elements

Periods	I	II	III	IV	V	VI	VII	VIII			0
1	1 H Hydrogen 1·0080										2 He Helium 4·003
2	3 Li Lithium 6·940	4 Be Beryllium 9·02	5 B Boron 10·82	6 C Carbon 12·010	7 N Nitrogen 14·008	8 O Oxygen 16·0000	9 F Fluorine 19·000				10 Ne Neon 20·183
3	11 Na Sodium 22·997	12 Mg Magnesium 24·32	13 Al Aluminium 26·97	14 Si Silicon 28·06	15 P Phosphorus 30·98	16 S Sulphur 32·06	17 Cl Chlorine 35·457				18 A Argon 39·944
4	19 K Potassium 39·096	20 Ca Calcium 40·08	21 Sc Scandium 45·10	22 Ti Titanium 47·90	23 V Vanadium 50·95	24 Cr Chromium 52·01	25 Mn Manganese 54·93	26 Fe Iron 55·85	27 Co Cobalt 58·94	28 Ni Nickel 58·69	
	29 Cu Copper 63·57	30 Zn Zinc 65·38	31 Ga Gallium 69·72	32 Ge Germanium 72·60	33 As Arsenic 74·91	34 Se Selenium 78·96	35 Br Bromine 79·916				36 Kr Krypton 83·7
5	37 Rb Rubidium 85·48	38 Sr Strontium 87·63	39 Y Yttrium 88·92	40 Zr Zirconium 91·22	41 Nb Niobium 92·91	42 Mo Molybdenum 95·95	43 Tc Technetium 99	44 Ru Ruthenium 101·7	45 Rh Rhodium 102·91	46 Pd Palladium 106·7	
	47 Ag Silver 107·880	48 Cd Cadmium 112·41	49 In Indium 114·76	50 Sn Tin 118·70	51 Sb Antimony 121·76	52 Te Tellurium 127·61	53 I Iodine 126·92				54 Xe Xenon 131·3
6	55 Cs Caesium 132·91	56 Ba Barium 137·36	57 La* Lanthanum 138·92	72 Hf Hafnium 178·6	73 Ta Tantalum 180·88	74 W Tungsten 183·92	75 Re Rhenium 186·31	76 Os Osmium 190·2	77 Ir Iridium 193·1	78 Pt Platinum 195·23	
	79 Au Gold 197·2	80 Hg Mercury 200·61	81 Tl Thallium 204·39	82 Pb Lead 207·21	83 Bi Bismuth 209·00	84 Po Polonium 210	85 At Astatine 211				86 Rn Radon 222
7	87 Fr Francium 223	88 Ra Radium 226·05	89 Ac** Actinium 227								

6	*58 – 71 Lanthanide series	58 Ce Cerium 140·13	59 Pr Praseodymium 140·92	60 Nd Neodymium 144·27	61 Pm Promethium 147	62 Sm Samarium 150·43	63 Eu Europium 152·0	64 Gd Gadolinium 156·9	65 Tb Terbium 159·2	66 Dy Dysprosium 162·46	67 Ho Holmium 164·94	68 Er Erbium 167·2	69 Tm Thulium 169·4	70 Yb Ytterbium 173·04	71 Lu Lutecium 174·99
7	**90 – 101 Actinide series	90 Th Thorium 232·12	91 Pa Protactinium 231	92 U Uranium 238·07	93 Np Neptunium 237	94 Pu Plutonium 239	95 Am Americium 241	96 Cm Curium 242	97 Bk Berkelium 243	98 Cf Californium 244	99 E Ein- steinium	100 Fm Fermium	101 Mv Mende- leevium		

Bibliography

1 T. A. AGEKYAN; Fluctuations in the apparent distribution of stars (O flyuktuatsiyakh v vidimykh raspredeleniyakh zvezd), *Astronomicheskiĭ Zhurnal* **32**, 416, 1955.

2 L. H. ALLER; *Astrophysics*, Vol. I, *The atmospheres of the Sun and stars*, Ronald Press Company 1953.

3 V. A. AMBARTSUMYAN; The excitation of the metastable states in the gaseous nebulae, *Tsirkulyary Glavnoĭ Astronomicheskoĭ Observatorii* No. 6, 10, 1933.

4 V. A. AMBARTSUMYAN; On the intensities of λ 4686 and H_β in the Wolf-Rayet stars, *ibid.* No. 7, 11, 1933.

5 V. A. AMBARTSUMYAN; On the radiative equilibrium of a planetary nebula, *Izvestiya Glavnoĭ Astronomicheskoĭ Observatorii* No. 114, 1933.

6 V. A. AMBARTSUMYAN; On the ionisation in the nebular envelope surrounding a star, *Monthly Not. R. Astr. Soc.* **95**, 469, 1935.

7 V. A. AMBARTSUMYAN; Compton scattering in stellar spectra (Komptonovskoe rasseyanie v spektrakh zvezd), *Uchenye Zapiski Leningradskogo Gosudarstvennogo Universiteta* No. 22, 5, 1938.

8 V. A. AMBARTSUMYAN; Radiative equilibrium of hydrogen envelopes round stars (O luchevom ravnovesii vodorodnykh obolochek, okruzhayushchikh zvezdy), *ibid.* No. 31, 1939.

9 V. A. AMBARTSUMYAN; Diffusion of light through scattering medium of large optical thickness, *Comptes Rendus de l'Acad. des Sci. de l'URSS* **43**, 102, 1944.

10 V. A. AMBARTSUMYAN; Stellar associations (Zvezdnye assotsiatsii), *Astronomicheskiĭ Zhurnal* **26**, 3, 1949; *Abh. sowj. Astr.* **1**, 33, 1951.

11 V. A. AMBARTSUMYAN and B. E. MARKARYAN; A stellar association near P Cygni (Zvezdnaya assotsiatsiya vokrug P Lebedya), *Soobshcheniya Byurakanskoĭ Observatorii* No. 2, 1949; *Abh. sowj. Astr.* **1**, 43, 1951.

12 R. G. ATHAY, D. E. BILLINGS, J. W. EVANS and W. O. ROBERTS; Emission in hydrogen Balmer lines and continuum in flash spectrum of 1952 total solar eclipse at Khartoum, Sudan, *Astrophys. J.* **120**, 94, 1954.

13 H. W. BABCOCK; The solar magnetograph, *Astrophys. J.* **118**, 387, 1953.

14 H. W. BABCOCK and T. G. COWLING; General magnetic fields in the Sun and stars, *Monthly Not. R. Astr. Soc.* **113**, 357, 1953.

15 J. G. BAKER and D. H. MENZEL; Physical processes in gaseous nebulae. III. The Balmer decrement, *Astrophys. J.* **88**, 52, 1938.

16 D. BARBIER, D. CHALONGE and N. MORGULEFF; Equivalent widths of lines of hydrogen and ionised calcium in stellar spectra (Intensités totales des raies de l'hydrogène et du calcium ionisé dans les spectres stellaires), *Annales d'Astrophysique* **4**, 137, 1941.

17 T. J. BARTLETT, B. WITTE and W. O. ROBERTS; Eruptive prominence associated with the limb flare of May 8, 1951, *Astrophys. J.* **117**, 292, 1953.

18 C. S. BEALS; Spectrophotometric studies of Wolf Rayet stars and novae, *Publ. Dominion Astrophys. Obs. Victoria* **6**, 93, 1934.

19 C. S. BEALS and R. D. HATCHER; The photoelectric temperatures of the P Cygni stars, *Contr. Dominion Obs. Ottawa* **2**, No. 1, 1948.

V. S. BERDICHEVSKAYA; The structure of the photospheres of A 0-type stars (Stroenie fotosfer zvezd tipa A 0), *Astronomicheskiĭ Zhurnal* **22**, 359, 1945.

21 V. S. Berdichevskaya; The structure of sunspots on the theory of radiative equilibrium (Stroenie solnechnykh pyaten po teorii luchistogo ravnovesiya), *ibid.* **31**, 51, 1954.

22 L. Biberman; On the theory of the diffusion of resonance radiation (K teorii diffuzii rezonansnogo izlucheniya), *Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki* **17**, 416, 1947.

23 A. F. Bogorodskiĭ and N. A. Khinkulova; The electron density distribution in the solar corona (O raspredelenii élektronnoĭ plotnosti v solnechnoĭ korone), *Publikatsii Kievskoĭ Astronomicheskoĭ Observatorii* No. **4**, 3, 1950.

24 M. Born; *Optik*, Springer, Berlin 1933.

25 E. Ya. Bugoslavskaya; The structure of the solar corona (Struktura solnechnoĭ korony), *Trudy Gosudarstvennogo Astronomicheskogo Instituta im. P. K. Shternberga* **19**, 1950.

26 I. W. Busbridge; Coherent and non-coherent scattering in the theory of line formation, *Monthly Not. R. Astr. Soc.* **113**, 52, 1953.

27 W. Buscombe; Spectrophotometry of early A-type stars, *Astrophys. J.* **114**, 73, 1951.

28 S. Chandrasekhar; *An introduction to the study of stellar structure*, University of Chicago Press 1939.

29 S. Chandrasekhar; On the radiative equilibrium of a stellar atmosphere. II, *Astrophys. J.* **100**, 76, 1944.

30 S. Chandrasekhar; On the radiative equilibrium of a stellar atmosphere. X, *ibid.* **103**, 351, 1946.

31 S. Chandrasekhar and F. H. Breen; On the continuous absorption coefficient of the negative hydrogen ion. III, *ibid.* **104**, 430, 1946.

32 G. G. Cillié and D. H. Menzel; The physical state of the solar chromosphere, *Harvard Circ.* No. 410, 1935.

33 W. J. Claas; The composition of the solar atmosphere, *Recherches Astronomiques de l'Observatoire d'Utrecht* **12**, Part 1, 1951.

34 T. G. Cowling; The electrical conductivity of an ionized gas in a magnetic field, with applications to the solar atmosphere and the ionosphere, *Proc. Roy. Soc.* A **183**, 453, 1945.

35 T. G. Cowling; The growth and decay of the sunspot magnetic field, *Monthly Not. R. Astr. Soc.* **106**, 218, 1946.

36 H. W. Dodson, E. R. Hedeman and L. Owren; Solar flares and associated 200 Mc/sec radiation, *Astrophys. J.* **118**, 169, 1953; H. W. Dodson, E. R. Hedeman and A. E. Covington; Solar flares and associated 2800 Mc/sec (10·7 cm) radiation, *ibid.* **119**, 541, 1954.

37 H. W. Dodson and R. R. McMath; The limb flare of May 8, 1951, *ibid.* **115**, 78, 1952.

38 V. A. Dombrovskiĭ; The polarisation of the light of early-type stars (O polyarizatsii izlucheniya zvezd rannikh spektral'nykh tipov), *Doklady Akademii Nauk Armyanskoĭ SSR* **10**, No. 5, 1949.

39 V. A. Dombrovskiĭ; A study of the polarisation of the light from stars near the stellar association Cepheus I (Izuchenie polyarizatsii sveta zvezd v oblasti zvezdnoĭ assotsiatsii Tsefeĭ I), *ibid.* **12**, No. 4, 1950.

40 M. A. Ellison; A photometric survey of solar flares, plages and prominences in H_α light, *Publ. Roy. Obs. Edinburgh* **1**, 75, 1952.

41 M. A. Ellison; A photometric survey of solar flares, plages and prominences in H_α light, carried out at the Royal Observatory, Edinburgh, 1947—52, *Atti dei Convegni dell' Accademia Nazionale dei Lincei* No. 11, 84, 1953.

42 T. M. Fofanova; Spectrophotometry of supergiant stars (Spektrofotometriya zvezd-sverkhgigantov), *Izvestiya Glavnoĭ Astronomicheskoĭ Observatorii* No. 144, 68, 1950.

43 V. F. Gaze and G. A. Shaĭn; An investigation of the spectrum of carbon stars (Issledovanie spektra uglerodnykh zvezd), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **2**, 51, 1948.

44 L. Goldberg, H. W. Dodson and E. A. Müller; The width of H_α in solar flares, *Astrophys. J.* **120**, 83, 1954.

45 V. G. Gorbatskiĭ; The interpretation of the changes in spectrum of γ Cassiopeiae (Interpretatsiya izmenenii spektra γ Kassiopei), *Astronomicheskii Zhurnal* **26**, 307, 1949.

46 V. G. Gorbatskiĭ; Ionisation in extended stellar envelopes (Ionizatsiya v protyazhennykh obolochkakh zvezd), *Uchenye Zapiski Leningradskogo Gosudarstvennogo Universiteta* No. 136, 72, 1950.

47 Sh. G. Gordeladze; The application of Zanstra's method to determine the temperatures of stars surrounded by a small gaseous envelope (Über die Anwendung der Zanstra-Methode zur Temperaturbestimmung der von einer Gashülle kleineren Umfangs umgebenen Sterne), *Zeitschrift für Astrophysik* **13**, 48, 1936.

48 Sh. G. Gordeladze; On the estimation of the mass ejected during the eruption of a nova (Akhal varskvlavta antebis dros gamotkortsnili airadi garsis shesakheb), *Byulleten' Abastumanskoĭ Astrofizicheskoĭ Observatorii* **1**, 55, 1937.

49 I. M. Gordon; The physical nature of chromospheric eruptions (K voprosu o fizicheskoĭ prirode khromosfernykh izverzheniĭ), *Doklady Akademii Nauk SSSR* **94**, 813, 1954.

50 J. L. Greenstein; Spectrophotometry of the F stars and of τ Ursae Majoris. I, *Astrophys. J.* **107**, 151, 1948; Analysis of metallic-line stars. II, *ibid.* **109**, 121, 1949.

51 J. L. Greenstein and E. Tandberg-Hanssen; The abundance of beryllium in the Sun, *ibid.* **119**, 113, 1954.

52 W. Grotrian; The physical interpretation of the light curve of Nova Herculis 1934 (Zur physikalischen Deutung der Lichtkurve der Nova Herculis 1934), *Zeitschrift für Astrophysik* **13**, 215, 1937.

53 L. É. Gurevich and A. I. Lebedinskiĭ; The magnetic field of sunspots (Magnitnoe pole solnechnykh pyaten) I, II, *Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki* **16**, 832, 840, 1946.

54 L. É. Gurevich and A. I. Lebedinskiĭ; Explosions in stars, resulting from nuclear reactions, as a possible cause of outbursts of novae and supernovae, *Comptes Rendus de l'Acad. des Sci. de l'URSS* **56**, 23, 1947; Peripheral explosions in stars as a result of nuclear reactions, *ibid.* **56**, 137, 1947.

55 L. É. Gurevich and A. I. Lebedinskiĭ; The theory of the chromosphere (Teoriya khromosfery), *Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki* **20**, 566, 1950.

56 J. S. Hall; Observations of the polarized light from stars, *Science* **109**, 166, 1949.

57 D. L. Harris III; On the line-absorption coefficient due to Doppler effect and damping, *Astrophys. J.* **108**, 112, 1948.

58 D. L. Harris III; Theoretical contours of absorption lines, *ibid.* **109**, 53, 1949.

59 G. Herzberg; *Atomic spectra and atomic structure*, Dover Publications, New York 1944 (tr. of *Atomspektren und Atomstruktur*, Steinkopff, Dresden 1936).

60 W. A. Hiltner; On the presence of polarization in the continuous radiation of stars. II, *Astrophys. J.* **109**, 471, 1949.

61 F. Hjerting; Tables facilitating the calculation of line absorption coefficients, *Astrophys. J.* **88**, 508, 1938.

62 G. S. Ivanov-Kholodnyi; A spectrophotometric investigation of hydrogen and helium in prominences (Spektrofotometricheskoe issledovanie vodoroda i geliya v protuberantsakh) I, *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **13**, 112, 1955.

63 E. JAHNKE and F. EMDE; *Tables of functions*, Dover Publications, New York 1945 (tr. of *Funktionentafeln*, 3rd ed., Teubner, Leipzig 1938).

64 J. T. JEFFERIES; Some electron collision cross sections of Ca II, *Australian J. Phys.* **7**, 22, 1954.

65 P. E. KOLPAKOV and YA. P. TERLETSKII; The problem of the propagation of the electromagnetic fields of sunspots through the Sun's atmosphere (K voprosu rasprostraneniya élektromagnitnogo polya pyaten v atmosfere Solntsa), *Doklady Akademii Nauk SSSR* **76**, 185, 1951.

66 T. V. KRAT; Continuous absorption in sunspot spectra (Nepreryvnoe pogloshchenie v spektre solnechnykh pyaten), *Izvestiya Glavnoĭ Astronomicheskoĭ Observatorii* No. 140, 56, 1947.

67 T. V. KRAT; The ionisation and electron pressure in solar faculae (Ionizatsiya i élektronnoe davlenie v solnechnykh fakelakh), *Astronomicheskiĭ Zhurnal* **24**, 329, 1947.

68 T. V. KRAT; The temperature and pressure in sunspots (Temperatura i davlenie v solnechnykh pyatnakh), *Izvestiya Glavnoĭ Astronomicheskoĭ Observatorii* No. 137, 1, 1948.

69 T. V. KRAT; The determination of line intensity gradients in the solar chromosphere (Opredelenie gradientov intensivnosti liniĭ solnechnoĭ khromosfery), *ibid.* No. 152, 20, 1954.

70 V. A. KRAT; The interpretation of spectroheliograms (Interpretatsiya spektrogeliogramm), *Uspekhi Astronomicheskikh Nauk* **3**, 87, 1947.

71 V. A. KRAT; On the positive charge of the Sun, *Comptes Rendus de l'Acad. des Sci. de l'URSS* **55**, 203, 1947.

72 V. A. KRAT; The spectra of the chromosphere and of faint prominences (O spektre khromosfery i slabykh protuberantsev), *Izvestiya Glavnoĭ Astronomicheskoĭ Observatorii* No. 147, 1, 1951.

73 V. A. KRAT; An investigation of granulation in the solar photosphere (Issledovanie granulyatsii solnechnoĭ fotosfery). I, *ibid.* No. 152, 1, 1954.

74 V. A. KRAT; The contour of the chromospheric line H_α (O konture khromosfernoĭ linii H_α), *Doklady Akademii Nauk SSSR* **97**, 45, 1954.

75 V. A. KRAT and T. V. KRAT; The contours of spectral lines in the solar chromosphere (O konturakh spektral'nykh liniĭ solnechnoĭ khromosfery), *Izvestiya Glavnoĭ Astronomicheskoĭ Observatorii* No. 153, 1, 1955.

76 V. A. KRAT and I. A. PROKOF'EVA; The structure of the solar chromosphere (O strukture solnechnoĭ khromosfery), *ibid.* No. 152, 9, 1954.

77 G. P. KUIPER; *The solar system*, Part I, *The Sun*, University of Chicago Press 1953.

78 L. LARMORE; A study of the motions of solar prominences, *Astrophys. J.* **118**, 436, 1953.

79 A. I. LEBEDINSKIĬ and S. S. ZHURAVLEV; Photoelectric photometry of the Sun's surface (Fotoélektricheskaya fotometriya solnechnoĭ poverkhnosti), *Astronomicheskiĭ Zhurnal* **18**, 312, 1941.

80 Liége, Mémoires de la Société Royale des Sciences de, 4me Série, Tome XIV: Les processus nucléaires dans les astres, 1954.

81 O. A. MEL'NIKOV; A new method of studying stellar spectra (Novyĭ metod izucheniya zvezdnykh spektrov), *Uspekhi Astronomicheskikh Nauk* **5**, 306, 1950.

82 O. A. MEL'NIKOV; Spectrophotometry of δ Cephei and η Aquilae and the K-effect for cepheids (Spektrofotometriya zvezd δ Tsefeya i η Orla i K-éffekt dlya tsefeid), *Trudy Glavnoĭ Astronomicheskoĭ Observatorii* Series II, No. 64, 1950.

83 O. A. MEL'NIKOV; Some results of spectrophotometry of A-type stars (Nekotorye rezul'taty spektrofotometrii zvezd tipa A), *Astronomicheskiĭ Zhurnal* **31**, 249, 1954.

84 D. H. Menzel, L. H. Aller and M. H. Hebb; Physical processes in gaseous nebulae, XIII. The electron temperatures of some typical planetary nebulae, *Astrophys. J.* **93**, 230, 1941.

85 D. H. Menzel and J. W. Evans; The behavior and classification of solar prominences, *Atti dei Convegni dell' Accademia Nazionale dei Lincei* No. 11, 119, 1953.

86 D. H. Menzel and C. L. Pekeris; Absorption coefficients and hydrogen line intensities, *Monthly Not. R. Astr. Soc.* **96**, 77, 1935.

87 R. Michard; The physical study of the photosphere and sunspots (Contribution à l'étude physique de la photosphère et des taches solaires), *Annales d'Astrophysique* **16**, 217, 1953.

88 M. Minnaert; The theoretical intensities of faint Fraunhofer lines, *Bull. Astr. Inst. Netherlands* **10**, 339, 399, 1948.

89 O. N. Mitropol'skaya; The construction of a curve of growth and an investigation of the physical conditions in solar faculae (Postroenie krivoĭ rosta i issledovanie fizicheskikh usloviĭ v solnechnykh fakelakh), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* 8, 93, 1952.

90 O. N. Mitropol'skaya; An investigation of the physical conditions in solar faculae (Issledovanie fizicheskikh usloviĭ v solnechnykh fakelakh), *ibid.* **11**, 152, 1954.

91 S. Miyamoto; Residual intensity and contour of the solar K-line, *Zeitschrift für Astrophysik* **31**, 282, 1953.

92 W. W. Morgan, P. C. Keenan and E. Kellman; *An atlas of stellar spectra, with an outline of spectral classification*, Astrophysical Monographs, University of Chicago 1943.

93 G. Münch; Model solar atmospheres, *Astrophys. J.* **106**, 217, 1947.

94 É. R. Mustel'; The theory of the radiative equilibrium of stellar atmospheres for an absorption coefficient depending on the frequency (Teoriya luchistogo ravnovesiya zvezdnykh atmosfer dlya koéffitsienta pogloshcheniya, zavisyashchego ot chastoty), *Trudy Gosudarstvennogo Astronomicheskogo Instituta im. P. K. Shternberga* **13**, No. 2, 5, 1940.

95 É. R. Mustel'; The distribution of energy in the continuous spectrum of stars of early spectral classes, *Astronomicheskiĭ Zhurnal* **18**, 297, 1941; **21**, 133, 1944.

96 É. R. Mustel'; On the expansion of new stars before maximum brightness, *ibid.* **22**, 65, 1945; The mechanism of the expansion of novae before maximum light and the masses of novae (O mekhanizme rasshireniya novykh zvezd pered maksimumom bleska i massy novykh zvezd), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **4**, 152, 1949.

97 É. R. Mustel'; The present scale of stellar temperatures (Sovremennaya shkala zvezdnykh temperatur), *Uspekhi Astronomicheskikh Nauk* **3**, 155, 1947.

98 É. R. Mustel'; The interpretation of the negative accelerations found from some absorption spectra of novae (Ob interpretatsii otritsatel'nykh uskoreniĭ, kharakterizovavshikh nekotorye absorbtsionnye sistemy v spektrakh novykh zvezd), I, *Astronomicheskiĭ Zhurnal* **24**, 280, 1947.

99 É. R. Mustel'; The origin of the diffuse-enhanced and Orion spectra in novae (O proiskhozhdenii diffuzno-iskrovogo i orionova spektra u novykh zvezd), *ibid.* **24**, 97, 155, 1947; An investigation of the ejection of matter by novae after maximum light (Issledovanie voprosa o vybrasyvanii materiĭ novymi zvezdami posle maksimuma bleska), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **1**, No. 2, 91, 1948.

100 É. R. Mustel'; The basic physical principles of the theory of radiative equilibrium in stellar atmospheres (Ob osnovnykh fizicheskikh printsipakh teorii luchevogo ravnovesiya zvezdnykh atmosfer), *ibid.* **2**, 123, 1948.

101 É. R. MUSTEL'; The interpretation of the negative accelerations found from some absorption spectra of novae (Ob interpretatsii otritsatel'nykh uskorenii, kharakterizovavshikh nekotorye absorbtsionnye sistemy v spektrakh novykh zvezd), II, *Astronomicheskii Zhurnal* **25**, 11, 1948.

102 É. R. MUSTEL'; The cause of nova outbursts (O proiskhozhdenii vspyshek novykh zvezd). *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **6**, 144, 1951.

103 É. R. MUSTEL'; The interpretation of calcium and hydrogen spectroheliograms (Ob istolkovanii kal'tsievykh i vodorodnykh spektrogeliogramm), *Doklady Akademii Nauk SSSR* **82**, 21, 1952.

104 É. R. MUSTEL'; The mechanism of the radiation of helium flocculi (O mekhanizme svecheniya gelievykh flokkulov), *ibid.* **97**, 625, 1954.

105 É. R. MUSTEL'; The mechanism of the radiation of hydrogen flocculi (O mekhanizme svecheniya vodorodnykh flokkulov), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **11**, 102, 1954.

106 É. R. MUSTEL'; The mechanism of the radiation of flocculi in the H_2 and K_2 lines (K voprosu o mekhanizme svecheniya flokkulov v liniyakh H_2 i K_2), *ibid.* **11**, 165, 1954.

107 É. R. MUSTEL'; The chromosphere above sunspots (O khromosfere nad solnechnymi pyatnami), *ibid.* **13**, 96, 1955.

108 É. R. MUSTEL'; The origin of the continuous spectrum of chromospheric flares (O proiskhozhdenii nepreryvnogo spektra khromosfernykh vspyshek), *ibid.* **15**, 54, 1955.

109 É. R. MUSTEL' and L. S. GALKIN; An investigation of stars of spectral classes A and F having anomalous metallic-line intensities (Issledovanie zvezd spektral'nykh klassov A i F s anomal'nymi intensivnostyami linii metallov), *ibid.* **12**, 148, 1954.

110 É. R. MUSTEL' and A. B. SEVERNYI; Spectrophotometry of chromospheric flares (Spektrofotometriya khromosfernykh vspyshek), *ibid.* **5**, 3, 1950.

111 É. R. MUSTEL' and A. B. SEVERNYI; Investigation of spectra of a large chromospheric flare on the Sun (Issledovanie spektrov bol'shoĭ khromosfernoĭ vspyshki na Solntse), *Doklady Akademii Nauk SSSR* 80, 867, 1951; *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* 8, 19, 1952.

112 L. NEVEN and C. DE JAGER; Observational models and helium abundances of the atmospheres of four B-type stars, *Bull. Astr. Inst. Netherlands* **12**, 103, 1954.

113 J. NEYMAN and E. L. SCOTT; A theory of the spatial distribution of galaxies, *Astrophys. J.* **116**, 144, 1952.

114 A. A. NIKITIN; An estimate of the abundance of helium in planetary nebulae (Otsenka obil'nosti geliya v planetarnykh tumannostyakh), *Uchenye Zapiski Leningradskogo Gosudarstvennogo Universiteta* No. 136, 47, 1950.

115 A. A. NIKITIN; The appearance of helium absorption lines in the Sun's spectrum (K voprosu o poyavlenii linii pogloshcheniya geliya v spektre Solntsa), *Astronomicheskii Zhurnal* **29**, 463, 1952.

116 V. B. NIKONOV and E. K. NIKONOVA; Absolute photoelectric photometry of the solar corona during the total eclipse of 9 July 1945 (Absolyutnaya élektrofotometriya solnechnoĭ korony vo vremya polnogo solnechnogo zatmeniya 9 iyulya 1945 g.), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **1**, No. 1, 83, 1947.

117 A. PANNEKOEK; The stellar temperature scale, *Astrophys. J.* 84, 481, 1936.

118 P. P. PARENAGO; The mass-luminosity relation, *Astronomicheskii Zhurnal* 14, 33, 1937.

119 P. P. PARENAGO; The structure of the Galaxy (Stroenie Galaktiki), *Uspekhi Astronomicheskikh Nauk* 4, 69, 1948; *Abh. sowj. Astr.* 3, 7, 1953.

120 P. P. PARENAGO and A. G. MASEVICH; Investigation of the mass-radius-luminosity relation (Issledovanie zavisimosti massa-radius-svetimost'), *Trudy Gosudarstvennogo Astronomicheskogo Instituta im. P. K. Shternberga* 20, 81, 130, 1951; *Abh. sowj. Astr.* 2, 27, 45, 1951.

121 J.-C. PECKER: On the theory of spectral types (Contribution à la théorie du type spectral), *Annales d'Astrophysique* **13**, 294, 319, 433, 1950.

122 W. PETRIE; Spectrophotometric gradients of some Wolf-Rayet, O, and B, stars, *Publ. Dominion Astrophys. Obs. Victoria* 7, 383, 1947.

123 J. H. PIDDINGTON; Model solar chromospheres, *Astrophys. J.* **119**, 531, 1954.

124 S. B. PIKEL'NER; The dissipation of the corona and its significance (Dissipatsiya korony i ee znachenie), *Doklady Akademii Nauk SSSR* **72**, 255, 1950.

125 R. S. RICHARDSON and M. SCHWARZSCHILD; On the turbulent velocities of solar granules, *Astrophys. J.* 111, 351, 1950.

126 K. ROTHSCHILD, J.-C. PECKER and W. O. ROBERTS; The prominence of July 25, 1951, *Astrophys. J.* 121, 224, 1955.

127 A. B. SEVERNYĬ; The turbulent state of the solar chromosphere (O turbulentnom sostoyanii solnechnoĭ khromosfery), *Doklady Akademii Nauk SSSR* 58, 1617, 1947.

128 A. B. SEVERNYĬ; An investigation of the light variations in solar prominences (Issledovanie kolebaniĭ v svechenii solnechnykh protuberantsev), *ibid.* **73**, 475, 1950.

129 A. B. SEVERNYĬ; An experiment in phosphor photography of solar prominences in infra-red light (Opyt fosforofotografirovaniya solnechnykh protuberantsev v infrakrasnykh luchakh), *ibid.* **76**, 189, 1951.

130 A. B. SEVERNYĬ; The chief types of motion in solar prominences (Osnovnye tipy dvizheniĭ v solnechnykh protuberantsakh), *ibid.* **82**, 25, 1952.

131 A. B. SEVERNYĬ; The total energy emitted by chromospheric flares (O polnoĭ énergii, izluchaemoĭ khromosfernymi vspyshkami), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* **9**, 3, 1952.

132 A. B. SEVERNYĬ; Magnetohydrodynamic phenomena on the surface of the Sun (O magnitogidrodinamicheskikh yavleniyakh u poverkhnosti Solntsa), *ibid.* 11, 129, 1954.

133 A. B. SEVERNYĬ and A. B. GIL'VARG; An interference-polarisation filter for studying the Sun, and an experiment in its use (Interferentsionno-polyarizatsionnyĭ fil'tr dlya issledovaniya Solntsa i opyt ego primeneniya), *ibid.* **4**, 3, 1949.

134 A. B. SEVERNYĬ and É. R. MUSTEL'; An investigation of the chromospheric flare of 13 June 1950 (Issledovanie khromosfernoĭ vspyshki 13 iyunya 1950 goda), *ibid.* **13**, 82, 1955.

135 A. B. SEVERNYĬ and E. F. SHAPOSHNIKOVA; An investigation of the development of chromospheric flares on the Sun (Issledovanie razvitiya khromosfernykh vspyshek na Solntse), *ibid.* **12**, 3, 1954.

136 G. A. SHAIN; Note on intensity decrement of Balmer lines in gaseous nebulae, *Tsirkulyary Glavnoĭ Astronomicheskoĭ Observatorii* No. 11, 8, 1934.

137 G. A. SHAIN; The emission lines of hydrogen in the spectra of long-period variables (Über die Emissionslinien des Wasserstoffs in den Spektren langperiodischer Veränderlicher), *Zeitschrift für Astrophysik* **10**, 73, 1935.

138 G. A. SHAIN; The occurrence of carbon isotopes in the spectra of N-type stars, *Byulleten' Abastumanskoĭ Astrofizicheskoĭ Observatorii* No. 6, 1, 1942; *Vestnik Akademii Nauk SSSR* 10, No. 10, 53, 1940; and later papers (with V. F. GAZE) in *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii*. See also *Abh. sowj. Astr.* 2, 7, 15, 1951; *Observatory* **64**, 255, 1942.

139 G. A. SHAIN; Some physical characteristics of the atmospheres of supergiants cB5—cA3, *Byulleten' Abastumanskoĭ Astrofizicheskoĭ Observatorii* No. 7, 83, 1943.

140 G. A. SHAIN; The nature of the emission spectrum of long-period variables (Priroda émissionnogo spektra dolgoperiodicheskikh peremennykh zvezd), *Izvestiya Akademii Nauk SSSR, Seriya Fizicheskaya* **9**, 161, 1945.

141 G. A. SHAĬN; The intensity of the emission lines in the corona spectrum (Ob intensivnosti émissionnykh liniĭ v spektre korony), *Doklady Akademii Nauk SSSR* 58, 767, 1947.

142 G. A. SHAĬN; Results of observations of the corona spectrum on 19 June 1936 (Rezul'taty nablyudeniya spektra korony 19 iyunya 1936 g.), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* 1, No. 1, 102, 1947.

143 G. A. SHAĬN; The cores of hydrogen lines in the spectra of early-type stars (O yadrakh vodorodnykh liniĭ v spektrakh zvezd rannikh klassov), *ibid.* 3, 46, 1948.

144 G. A. SHAĬN; Central intensities of strong absorption lines in stellar spectra (Tsentral'nye intensivnosti sil'nykh liniĭ pogloshcheniya v spektrakh zvezd), *ibid.* 7, 67, 1951.

145 G. A. SHAĬN and V. F. GAZE; The predominant orientation of filaments and other structural features in gaseous nebulae (O preimushchestvennoĭ orientatsii volokon i drugikh strukturnykh detaleĭ v gazovykh tumannostyakh), *Doklady Akademii Nauk SSSR* 82, 857, 1952.

146 G. A. SHAĬN and P. F. SHAĬN; Systematic differential line shifts in the spectra of white supergiants (Differentsial'nye sistematicheskie smeshcheniya liniĭ v spektrakh belykh sverkhgigantov), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* 4, 49, 1949.

147 G. A. SHAĬN and O. STRUVE; On the rotation of the stars, *Monthly Not. R. Astr. Soc.* 89, 222, 1929.

148 I. S. SHKLOVSKIĬ; The anomalous density distribution in the chromosphere (K voprosu ob anomal'nom raspredelenii plotnosteĭ v khromosfere), *Doklady Akademii Nauk SSSR* 68, 829, 1949.

149 I. S. SHKLOVSKIĬ; The ultra-violet radiation of the corona and chromosphere and the ionisation of the Earth's atmosphere (Ul'trafioletovoe izluchenie korony i khromosfery i ionizatsiya zemnoĭ atmosfery), *Izvestiya Krymskoĭ Astrofizicheskoĭ Observatorii* 4, 80, 1949; *Abh. sowj. Astr.* 3, 247, 1953.

150 G. F. SITNIK; The nature of sunspots (K voprosu o prirode solnechnykh pyaten), *Trudy Gosudarstvennogo Astronomicheskogo Instituta im. P. K. Shternberga* 11, No. 2, 5, 1939.

151 G. F. SITNIK; Hydrodynamic and thermodynamic theories of sunspots (Gidrodinamicheskie i termodinamicheskie teorii solnechnykh pyaten), *Uspekhi Astronomicheskikh Nauk* 1, 107, 1939.

152 V. V. SOBOLEV; The determination of the electron temperatures of the planetary nebulae and a refinement of the "nebulium" method for determining the temperatures of their nuclei (Opredelenie élektronnykh temperaturakh planetarnykh tumannosteĭ i utochnenie metoda "nebuliya" dlya opredeleniya temperatury ikh yader), *Uchenye Zapiski Leningradskogo Gosudarstvennogo Universiteta* No. 82, 3, 1941.

153 V. V. SOBOLEV; Radiation pressure in an expanding nebula (Svetovoe davlenie v rasshiryayushcheĭsya tumannosti), *Astronomicheskiĭ Zhurnal* 21, 143, 1944.

154 V. V. SOBOLEV; The intensity of emission lines in the spectra of early-type stars (Ob intensivnosti émissionnykh liniĭ v spektrakh zvezd rannikh klassov), *ibid.* 24, 13, 1947.

155 V. V. SOBOLEV; Non-coherent scattering of light in stellar atmospheres (Nekogerentnoe rasseyanie sveta v zvezdnykh atmosferakh), *ibid.* 26, 129, 1949.

156 V. V. SOBOLEV; The polarisation of scattered light (O polyarizatsii rasseyannogo sveta), *Uchenye Zapiski Leningradskogo Gosudarstvennogo Universiteta* No. 116, 3, 1949.

157 V. V. SOBOLEV; The scattering of light in the atmospheres of the Earth and planets (O rasseyanii sveta v atmosferakh Zemli i planet), *ibid.* No. 116, 17, 1949.

158 V. V. SOBOLEV; The radiation of stellar envelopes in the absence of radiative equilibrium (Svechenie zvezdnykh obolochek pri otsutstvii luchevogo ravnovesiya), *Astronomicheskiĭ Zhurnal* **27**, 81, 1950.

159 V. V. SOBOLEV; The theory of the radiation of novae (K teorii svecheniya novykh zvezd), *ibid.* **31**, 15, 1954.

160 V. V. SOBOLEV; The formation of absorption lines with non-coherent scattering of light (Obrazovanie liniĭ pogloshcheniya pri nekogerentnom rasseyanii sveta), *ibid.* **31**, 231, 1954.

161 V. V. SOBOLEV; Diffusion of radiation with redistribution in frequency (Diffuziya izlucheniya s pereraspredeleniem po chastotam) I, *Vestnik Leningradskogo Gosudarstvennogo Universiteta* No. 5, 85, 1955.

162 L. SPITZER, JR., and J. L. GREENSTEIN; Continuous emission from planetary nebulae, *Astrophys. J.* **114**, 407, 1951.

163 V. E. STEPANOV; A new law in the appearance of sunspots (O novoĭ zakonomernosti v vozniknovenii solnechnykh pyaten), *Astronomicheskiĭ Zhurnal* **25**, 209, 1948.

164 B. STRÖMGREN; The influence of electron captures on the contours of Fraunhofer lines, *Zeitschrift für Astrophysik* **10**, 237, 1935.

165 P. TEN BRUGGENCATE, H. GOLLNOW and F. W. JÄGER; An attempt to determine the true intensity distribution at the extreme limb of the Sun (Über einen Versuch, die wahre Intensitätsverteilung am äußersten Sonnenrand zu ermitteln), *Zeitschrift für Astrophysik* **27**, 223, 1950.

166 M. TUBERG; The variations of absorption-line contours across the solar disc, *Astrophys. J.* **103**, 145, 1946.

167 A. B. UNDERHILL; A model atmosphere for an early O-type star, *Publ. Dominion Astrophys. Obs. Victoria* 8, 357, 1950: On the luminosity effects of the hydrogen lines in O-type spectra, *ibid.* 8, 385, 1950.

168 A. UNSÖLD; *Physik der Sternatmosphären*, 2nd ed., Springer, Berlin 1955.

169 E. F. M. VAN DER HELD; The intensity and natural width of spectral lines (Intensität und natürliche Breite von Spektrallinien), *Zeitschrift für Physik* **70**, 508, 1931.

170 S. E. A. VAN DIJKE; A comparative study of the spectra of α Bootis and 70 Ophiuchi A, *Astrophys. J.* **104**, 27, 1946.

171 S. VERWEY; The Stark effect of hydrogen in stellar spectra, *Publ. Astr. Inst. Univ. Amsterdam* No. 5, 1936.

172 B. A. VORONTSOV-VEL'YAMINOV; The intensities and photometric effects of bright bands in the spectra of six Wolf-Rayet stars, and their temperatures (Intensivnosti i fotometricheskiĭ éffekt yarkikh polos v spektrakh shesti zvezd Vol'f-Raĭe i ikh temperatura), *Astronomicheskiĭ Zhurnal* **23**, 1, 1946.

173 B. A. VORONTSOV-VEL'YAMINOV; The white-blue sequence in the Russell diagram (Belo-golubaya posledovatel'nost' na diagramme Ressela), *ibid.* **24**, 83, 1947; The interpretation of the white-blue sequence in the Russell diagram (Interpretatsiya belo-goluboĭ posledovatel'nosti na diagramme Ressela), *ibid.* **24**, 145, 1947.

174 S. K. VSEKHSVYATSKIĬ and E. YA. BUGOSLAVSKAYA; Rotation of solar corona as determined from apparent displacements in its structure, *Comptes Rendus de l'Acad. des Sci. de l'URSS* **25**, 362, 1939; The inner corona on June 19, 1936. Investigation of coronal structure and changes occurring in the corona and chromosphere, *ibid.* **25**, 367, 1939.

175 V. P. VYAZANITSYN; A spectrophotometric investigation of solar prominences (Spektrofotometricheskoe issledovanie solnechnykh protuberantsev), *Izvestiya Glavnoĭ Astronomicheskoĭ Observatorii* No. 136, 1, 1947.

176 V. P. VYAZANITSYN; Self-absorption in prominences in the H and K lines of Ca^+ (Samopogloshchenie protuberantsev v Ca^+ liniyakh H i K), *ibid.* No. 139, 31, 1947.

177 V. P. VYAZANITSYN; Absolute photometry of the solar chromosphere (Absolyutnaya fotometriya solnechoĭ khromosfery), *ibid.* No. 147, 19, 1951.

178 M. WALDMEIER; The visibility function of solar faculae (Die Sichtbarkeitsfunktion der Sonnenfackeln), *Zeitschrift für Astrophysik* **26**, 147, 1949.

179 M. WALDMEIER; *Ergebnisse und Probleme der Sonnenforschung*, 2nd ed., Akademische Verlagsgesellschaft, Leipzig 1955.

180 J. W. WARWICK; Heights of solar flares, *Astrophys. J.* **121**, 376, 1955.

181 H. ZANSTRA; Investigations on planetary nebulae (Untersuchungen über planetarische Nebel), I, *Zeitschrift für Astrophysik* **2**, 1, 1930.

182 H. ZANSTRA; The value of observations of continuous spectra in the chromosphere and prominences, *Circ. Astr. Inst. Univ. Amsterdam* No. 1, 1950; *see also* S. G. M. HAUG; The visual continuous spectrum and the thickness of a prominence, *ibid.* No. 5a, 1952; D. KOELBLOED; On the thickness of prominences in the line of sight, *ibid.* No. 5b, 1952; R. O. REDMAN and H. ZANSTRA; The continuous spectrum of a prominence observed at the total solar eclipse of 1952 February 25, *ibid.* No. 6, 1952.

183 H. ZANSTRA; On scattering with redistribution and radiation pressure in a stationary nebula, *Bull. Astr. Inst. Netherlands* **11**, 1, 1949.

INDEX OF STARS

INDEX OF SYMBOLS

INDEX OF NAMES

INDEX OF SUBJECTS